ARI Research Note 96-45

Building and Retaining the Career Force: New Procedures for Accessing and Assigning Army Enlisted Personnel

Annual Report, 1993 Fiscal Year

John P. Campbell and Lola M. Zook
Human Resources Research Organization

Selection and Assignment Research Unit
Michael G. Rumsey, Chief

March 1996

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Human Resources Research Organization

Technical review by

Peter J. Legree
Michael G. Rumsey

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1996, March	2. REPORT TYPE Interim	3. DATES COVERED (from. . . to) October 1992-September 1993
4. TITLE AND SUBTITLE Building and Retaining the Career Force: New Procedures for Accessing and Assigning Army Enlisted Personnel--Annual Report, 1993 Fiscal Year		5a. CONTRACT OR GRANT NUMBER MDA903-89-C-0202 5b. PROGRAM ELEMENT NUMBER 0603007A
6. AUTHOR(S) John P. Campbell and Lola M. Zook, Editors (HumRRO)		5c. PROJECT NUMBER A792 5d. TASK NUMBER 2208 5e. WORK UNIT NUMBER C01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization (HumRRO) 66 Canal Center Plaza, Suite 400 Alexandria, VA 22314		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RS 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 96-45

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Prepared under Project Building the Career Force (Human Resources Research Organization, American Institutes for Research, Personnel Decisions Research Institute, U.S. Army Research Institute) COR: Michael Rumsey

14. ABSTRACT *(Maximum 200 words)*:

The Career Force research project is the second phase of an Army program to develop a selection and classification system for enlisted personnel, based on expected future performance. In the first phase, Project A, a large and versatile database was collected from a representative sample of Military Occupational Specialties (MOS) and used to (1) validate the Armed Services Vocational Aptitude Battery (ASVAB) and (2) develop and validate new predictor and criterion measures representing the domain of potential measures. Building on this foundation, Career Force research is finishing developing the selection/classification system and evaluating its effectiveness, with emphasis on assessing second-tour performance. This fourth year of the project completed analyses of test results from the Longitudinal Validation cohort, including the second-tour sample. Development of optimal test batteries for predicting first- and second-tour performance, attrition, and reenlistment prospects is continuing.

15. SUBJECT TERMS

Career Force	Criterion measures	Longitudinal validation	Personnel classification
personnel selection	Predictor measures	Project A	Second-tour performance

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	348	

FOREWORD

This document is a description of the research activities conducted during the fourth year of the Building the Career Force project. This project is the second phase of a research program of unprecedented scope and depth to provide the basis for improving the Army's selection and classification procedures, as well as improving reenlistment and promotion decisions for soldiers up to the level of sergeant. The thrust for this program came from the practical, professional, and legal need to validate the Armed Services Vocational Aptitude Battery (ASVAB), the current U.S. military selection/classification test battery, and other selection variables as predictors of training and performance. The authorization for the program was provided in a letter, Deputy Chief of Staff for Operations, "Army Research Project to Validate the Predictive Value of the Armed Services Vocational Aptitude Batter," effective 19 November 1980, and a memorandum, Assistant Secretary of Defense, Manpower Reserve Affairs and Logistics (MRA&L), "Enlistment Standards," effective 11 September 1980.

The research program began in 1982 with an effort known as Project A. Project A not only validated the ASVAB against job performance, it further linked indicators of temperament (achievement, discipline, stress tolerance), psychomotor ability (e.g., eye-hand coordination), and spatial ability to job performance. Project A developed new tools for a variety of personnel decisions. Before these tools can be optimally used, however, two critical questions need to be answered: (1) What combinations of aptitude, temperament, psychomotor ability, and spatial ability, measured at or before entry into the Army, best predict later performance in individual military occupational specialties (MOS)? and (2) Which indicators of first-tour performance best predict performance in the second tour? These questions will be answered in Building the Career Force.

The fourth-year Building the Career Force activities described herein continued analyses focused on the combined set of initial entry predictor measures developed for selection and classification purposes, and end-of-training and first-tour job performance measures to be linked to these predictor measures. Data from second-tour measures administered to a sample already tested on initial entry, end-of-training, and first-tour measures are being examined for longitudinal linkages across the full set of measures from initial entry into second tour. The information obtained from these analyses will provide a base for setting selection, classification, reenlistment, and promotion policies unrivaled anywhere.

Active sponsorship of this effort has been provided by the Director of Military Personnel Management (DMPM). The DMPM has been periodically briefed on the activities described herein and has personally taken part in the execution of this project in an extremely effective manner. To ensure that Building the Career Force research

achieves its full scientific potential, an advisory group composed of experts in personnel measurement, selection, and classification has provided continuing guidance on technical aspects of the research. Members of this scientific advisory group include Drs. Philip Bobko, Lloyd Bond, Milton Hakel (Chair), Lloyd Humphreys, Lawrence Johnson, Robert Linn, Mary Tenopyr, and Jay Uhlaner.

ZITA M. SIMUTIS
Deputy Director
(Science and Technology)

EDGAR M. JOHNSON
Director

EDITORS' PREFACE

This is the fourth annual report for work completed as part of the Building the Career Force Project. It also constitutes the primary technical report of the work completed on several of the project's principal tasks. Consequently, it is a "stand alone" document for Fiscal Year 1993 and does not refer the reader to more detailed descriptions in supplementary reports for that period. The Career Force Project extends the major work on selection and classification of Army enlisted personnel that was completed as part of Project A.

The Career Force Project includes (1) a replication and extension of the Experimental Predictor Battery validities for the selection and classification of first-tour enlisted personnel; (2) validation of the Experimental Battery against end-of-training performance; (3) validation of training performance as a predictor of first-tour job performance; (4) measurement of second-tour performance; (5) validation of the Armed Services Vocational Aptitude Battery (ASVAB), the Experimental Battery, Advanced Individual Training (AIT) performance, and first-tour performance as predictors of second-tour performance; and (6) identification of the optimal predictor battery for selection and classification, given certain specific sets of goals and constraints.

The annual report for year one described the results of a series of analyses directed at basic score development for (1) the Experimental Predictor Battery, (2) the End-of-Training performance measures, and (3) the second-tour job performance measures that were administered to the second-tour Concurrent Validation sample (CVII). The performance data from this initial sample of second-tour junior noncommissioned officers (NCO) were also used to develop a latent structure model of second-tour performance. The model hypothesizes six basic components for NCO performance.

The annual report for year two dealt with the analysis of performance data from the Longitudinal Validation I (LVI) sample, which is a sample of approximately 10,000 first-tour incumbents who entered the Army during 1986/87. It is the second of the two major cohorts of enlisted personnel that make up the total Project A/Career Force Project data base. The criterion score development, data editing, and performance modeling analyses were each described in turn. The remainder of the report described the results of the basic Longitudinal Validation of the ASVAB and the Project A Experimental Predictor Battery against (1) training performance, (2) first-tour job performance, and (3) second-tour job performance (i.e., the second-tour performance factor scores developed during year one).

The third annual report, for FY92, focused exclusively on the Longitudinal Validation Second-Tour (LVII) sample. This is a sample of approximately 1,500 individuals in the nine "Batch A" MOS who had reenlisted and were 2-3 years into their second tour of duty at the time the LVII measures of second-tour job performance were administered. The individuals in the sample had entered the Army in 1986/87 as part of the Project A Longitudinal Validation (LV) sample. The report describes (1) the data collection procedure, (2) the editing of the data file, (3) the initial analyses of each instrument to develop the basic criterion scores, and (4) the development of a model of second-tour performance based on the LVII sample

data. A major feature of the results is the great consistency in the covariance structure of the basic criterion scores across cohorts (CVII vs. LVII) and across organizational levels (LVI vs. LVII).

This fourth annual report, for FY93, covers (1) a summary of the analyses done with the Experimental Battery to support the work of the Manpower Accession Policy Working Group as it considered future revisions to the ASVAB; (2) the basic validation analyses of ASVAB and ABLE against second-tour performance in the longitudinal sample (LVII), using the LVII performance model factors as criterion scores; (3) the degree to which correlations of performance with performance exhibit convergent and divergent validity across organizational levels (e.g., first-tour performance vs. second-tour performance) relative to the individual components of performance; (4) the results of modeling the predictors of first-tour attrition using event-history analysis; and (5) a description of results obtained with the Army Job Satisfaction Questionnaire.

The remaining activites for the project are to (1) identify the "optimal" prediction equations that maximize overall selection validity for the Army personnel system, given constraints; (2) estimate the potential differential prediction/classification validity of the Experimental Battery; (3) analyze the relative predictability of alternative selection and classification goals; and (4) conduct a complete fairness analysis of all major predictor/criterion relationships in the Project A/Career Force data base. The results of these analyses will be the topics of subsequent reports.

As was the case for previous years, the writing of this report was very much a collaborative effort by a lot of people. The primary authors for each chapter are indicated in the Contents and also on the first page of each chapter. The editors, and the management, are deeply appreciative of their contributions.

BUILDING AND RETAINING THE CAREER FORCE: NEW PROCEDURES FOR ACCESSING AND ASSIGNING ARMY ENLISTED PERSONNEL -- ANNUAL REPORT, 1993 FISCAL YEAR

EXECUTIVE SUMMARY

Requirements:

The Career Force Project is the second phase of a comprehensive, long-term research program, sponsored by the Deputy Chief of Staff for Personnel, to improve the selection and assignment of Army enlisted personnel. In the first phase, Project A, existing selection measures were validated against both existing and newly developed performance criteria, and new predictive measures were developed to aid in assignment and promotion decisions. The Career Force project extends the research to measure second-tour job performance, and to examine how selection and classification tests administered before a soldier's enlistment can, with measures of performance during that enlistment, predict performance potential for second-tour duty.

Procedure:

In Task 1, measures adopted in Project A to assess the performance of second-tour soldiers were administered to the Longitudinal Validation (LV) sample first tested in Project A at entry and during their first tour. The results from these tests are being analyzed to complete the predictive validation of the Armed Services Vocational Aptitude Battery (ASVAB) and the Project A Experimental Predictor Battery, measures of training success, and first-tour job performance tests against the criteria of successful second-tour performance. Task 2 staff has established an integrated data base and is processing Project A and Career Force data and merging files with related military data. Task 3 covers all analyses performed to develop the analytic framework needed to evaluate equations for predicting training performance, first-tour performance and attrition, reenlistment, and second-tour performance.

Findings:

The pattern of results from confirmatory analyses of the latest Longitudinal Validation tests has continued to be consistent with the results from earlier LV testing, as well as from the initial Concurrent Validation tests. The description of the latent structure of performance as individuals move from training through their first tour and into their second tour continued to be highly consistent as alternative ways of predicting and assessing development and leadership qualities are tested.

Performance in training does predict performance on the job (first tour), and entry-level performance does predict second-tour NCO performance.

The validity of the ASVAB is as high for predicting performance by soldiers in their second tour as it was for predicting first-tour performance.

Its highest validities are for predicting Core Technical Proficiency but it also has high validity for predicting second-term leadership performance.

Components of the Predictor Test Battery and measures of current performance each contribute considerable unique variance to predicting future performance.

Extensive analyses comparing various combinations of Experimental Battery tests and ASVAB subtests show that all relevant indices (e.g., absolute validity, discriminant validity, subgroup differences) cannot be maximized by the same test battery. In choosing tests to reconstitute a test battery, trade-offs will have to be made in terms of the objectives that are considered most important.

A model of the relationship of the ASVAB and Experimental Battery predictors with attrition shows that, as has long been recognized, high school diploma graduate status is the best predictor of low attrition among soldiers in their first tour. For incremental predictive gain, the best fitting predictor composite in these analyses consisted of five scales (e.g., Nondelinquency, Physical Condition) from the Assessment of Background and Life Experiences (ABLE) and the Quantitative subtest composite from the ASVAB.

Utilization of Findings:

The data from the completed predictor and performance analyses from Project A/Career Force are providing the base for a series of maximizing/optimizing analyses of selection and classification efficiency. These results will be integrated with previous validation and modeling work to develop comprehensive procedures for estimating potential gains to the Army from classification changes under varied psychometric, organizational, and labor market conditions.

BUILDING AND RETAINING THE CAREER FORCE: NEW PROCEDURES FOR ACCESSING AND ASSIGNING ARMY ENLISTED PERSONNEL -- ANNUAL REPORT, 1993 FISCAL YEAR

CONTENTS

CONTENTS (Continued)

List of Tables

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

Page

List of Tables (Continued)

List of Figures

Page

List of Figures (Continued)

BUILDING AND RETAINING THE CAREER FORCE: NEW PROCEDURES FOR ACCESSING AND ASSIGNING ARMY ENLISTED PERSONNEL -- ANNUAL REPORT, 1993 FISCAL YEAR

Chapter 1
INTRODUCTION

John P. Campbell and James H. Harris

This report is a summary of the major activities undertaken during the fourth year of a Department of the Army research project entitled Building and Retaining the Career Force. The report covers the period of the 1993 fiscal year, beginning 1 October 1992. The research reported was conducted by a consortium comprised of Human Resources Research Organization (HumRRO), American Institutes for Research (AIR), and Personnel Decisions Research Institute, Incorporated (PDRI, Inc.), under contract to and in collaboration with the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI).

The research effort is the second phase of a two-phase program to develop a selection and classification system for enlisted personnel, based on expected future performance. Phase One was Project A (Campbell & Zook, 1991). Its goals were to validate the Armed Services Vocational Aptitude Battery (ASVAB) by collecting data from a representative sample of Military Occupational Specialties (MOS), and to build a large and versatile data base by developing and validating new predictors and criterion measures that represented the entire domain of potential measures.

The goals of Building the Career Force are to determine the longitudinal relationship between the new predictors and first-tour performance, to finalize and administer the measures of second-tour job performance, and to examine how selection and classification tests administered before a soldier's first enlistment, in conjunction with performance during that soldier's first enlistment, predict performance in a second enlistment.

The remainder of this chapter describes the objectives and organization of the project, summarizes the work completed during the first years of the Project, and outlines the content included in this fourth annual report.

BUILDING THE CAREER FORCE: OBJECTIVES AND PROJECT DESIGN

The Project A data base, the predictor and criterion measures the project developed, the working models it provided, and its basic analytic work have provided a valuable foundation for the further production of scientific findings and operational products, and for the subsequent investigation of reenlistment decisions, noncommissioned officer (NCO) job performance, NCO promotion decisions, and the identification of NCO potential.

The work encompassed by the Career Force Project is intended to accomplish several general goals relevant to building and retaining the career force. The goals may be summarized as follows.

(1) Build the final pieces required for a complete selection/ classification decision-making system for Army enlisted personnel.

(2) Provide the analytic procedures and data necessary to maximize the system's performance and evaluate its effectiveness.

(3) Build the foundation for its implementation.

The principal focus is on the greatest possible gains in overall individual performance, for both "can do" and "will do" components of performance, that can be obtained from enhancing the selection/classification system for first- and second-tour enlisted personnel. Maximizing the benefit from a more effective match of people and jobs has always been a goal of the Army personnel system. Given the population demographics in prospect for the United States during the current decade, this goal becomes even more crucial. It is incumbent on virtually every organization to go as far as the state-of-the-art will allow.

This means that the information used to make personnel decisions must yield the maximum gain in terms of <u>accuracy</u> and <u>fairness</u> of predictions. It means that the models and procedures used to execute selection and classification decisions (e.g., new computerized tests) must both serve the goals of the organization and maximize the <u>aggregate</u> <u>benefits</u> that can be obtained from using the selection/ classification measures. It means that the implementation of the system, or any part of it, must serve the needs of the users and also maintain fidelity with the goals on which the system is based.

Specific Research Objectives

The specific scientific objectives of Building the Career Force are to:

(1) Develop a complete array of valid and reliable measures of second-tour performance as an Army NCO, using the Project A prototypes as a starting point.

(2) Carry out a complete incremental predictive validation of (a) the ASVAB and the Project A Experimental Battery of predictors, (b) measures of training success, and (c) the full array of first-tour performance criteria developed as part of Project A. The criteria against which these three sets of predictors will be validated, both individually and incrementally for each major criterion component, are the second-tour job performance measures.

(3) Develop a model of second-tour soldier performance that parallels the first-tour performance model from Project A and that identifies the major components of second-tour performance, provides information on their construct validity, and establishes how the components should be combined for specific prediction or interpretation purposes.

(4) Develop the analytic framework needed to evaluate the optimal prediction equations for predicting (a) training performance; (b) first-tour performance; (c) first-tour attrition and the reenlistment decision; and (d) second-tour performance, under the

conditions when testing time is limited to a specified amount and when there must be a tradeoff among alternative selection/classification goals (e.g., maximizing aggregate performance vs. minimizing discipline and low-motivation problems vs. minimizing attrition).

(5) Design and develop a fully functional and user-friendly research data base that includes all relevant personnel data on 1981/82, 1983/84, and 1986/87 accessions, including all Project A and Career Force Project data and all relevant data from the Enlisted Master File (EMF), Accession File, and Army Training Requirements and Resources System (ATRRS).

Project Organization

To reflect the requirements of the Career Force research, the project is organized as shown in Figure 1.1. Management of the total project is the responsibility of the Project Director. The overall design, execution, and evaluation of the substantive tasks are the responsibility of the Principal Scientist. Oversight and scientific participation are provided by the Army Research Institute. Guidance has been provided by the General Officers Steering Committee and the Scientific Advisory Group.

A brief summary of the work encompassed by the three substantive technical tasks follows:

Task 1 deals with revision of the measures developed in Project A to measure second-tour soldier performance. The second-tour performance measures were revised and were administered to the Project A Longitudinal Validation (LV) sample, beginning in June 1991. At that time, the soldiers in the sample were in their second tour, and had been in the Army anywhere from 41 to 63 months. Once the data have been fully analyzed (under Task 3), it will be possible to complete the incremental predictive validation of the ASVAB and the Project A Experimental Battery, the measures of training success, and the full array of first-tour performance measures developed in Project A, against the second-tour criterion measures.

Task 2 has a single purpose--to establish, manage, and safeguard an integrated research data base (IRDB). As part of the establishment of the IRDB, Task 2 is integrating the Project A longitudinal research data base, extracting and merging data from other military data bases, processing data collected by Project A and this project, and creating workfiles for analyses.

Task 3 is responsible for all analyses performed under this project. The task is organized around the five major data sets to be analyzed: the Longitudinal Validation predictor data (LV), the Longitudinal Validation end-of-training (EOT) data, the Longitudinal Validation first-tour data (LVI), the Concurrent Validation second-tour data (CVII), and the Longitudinal Validation second-tour data (LVII). At the end of the project, Task 3 will have developed the analytic framework necessary to evaluate optimal prediction equations to predict training performance, first-tour performance and attrition, reenlistment, and second-tour performance.

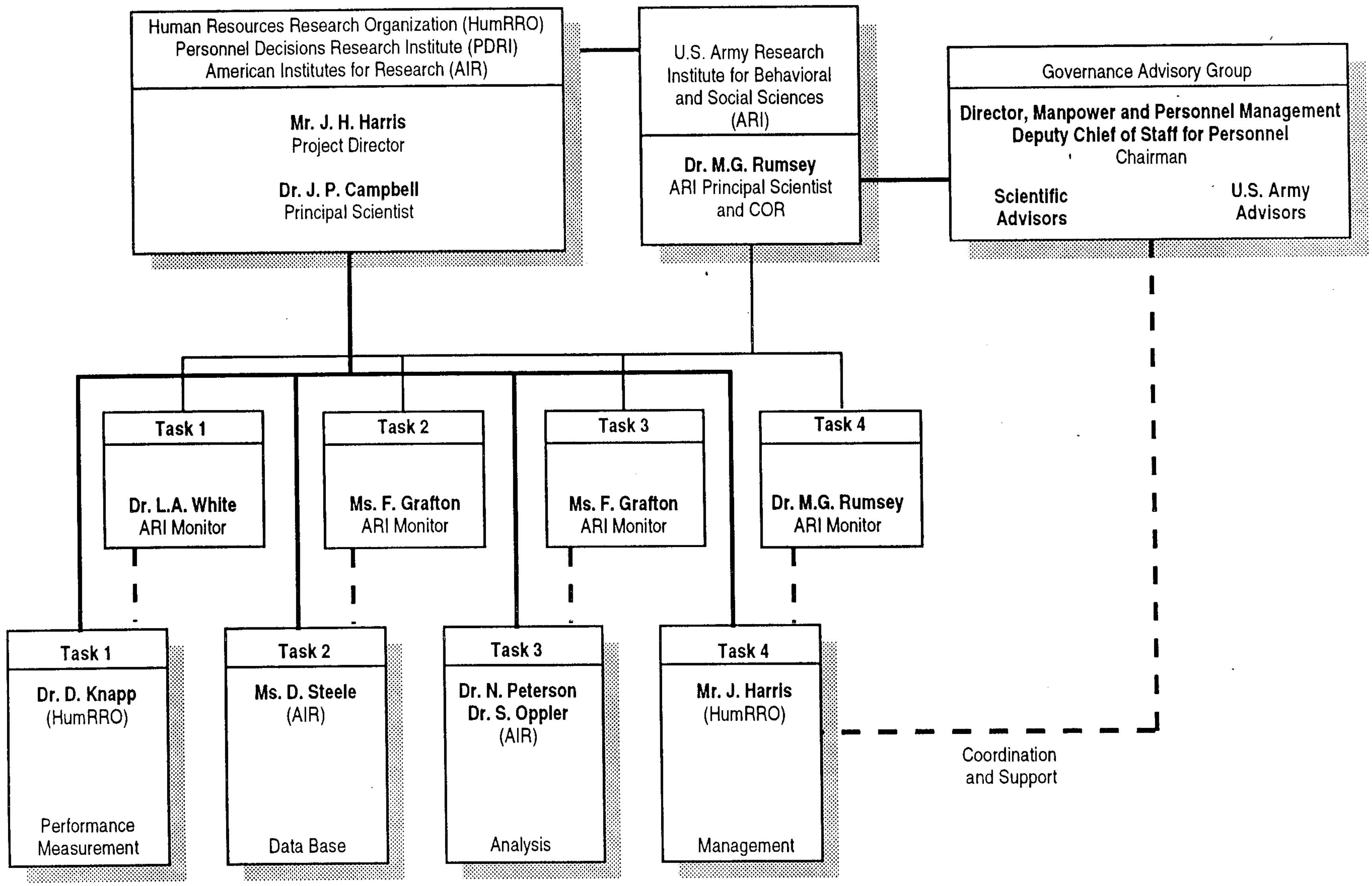


Figure 1.1. Building the Career Force: Project management structure.

Project Design

As will be explained in later sections of this chapter, the remaining chapters of this report all deal with the analyses of data obtained at one major point in the total project design. To set the stage for these discussions, as well as for the summary of work done during years one through three, the basic overall project design is summarized below.

The Research Sample

In general, the combined design for Project A/Career Force encompasses two major cohorts of soldiers (new accessions for 1983/84 and for 1986/87), each of which were followed into their second tour of duty and which collectively have produced six major research samples. For each research sample there is a battery of predictor measures and an array of performance measures. For each of the six samples the predictor battery is composed of the ASVAB and either the Trial Battery or the Experimental Battery version of the new tests developed in Project A (see Campbell & Zook, 1991). There were three distinct arrays of performance measures corresponding to the need to assess (a) training performance, (b) first-tour job performance, and (c) second-tour job performance.

In each sample the individuals to be assessed were selected from two predetermined sets of MOS -- Batch A and Batch Z. They are listed in Figure 1.2. The two groups differed in that tests administered to Batch A MOS included MOS-specific rating scales and job knowledge and hands-on tests, whereas the only MOS-specific measure administered to the Batch Z MOS was an end-of-training test.

Batch A		Batch Z	
MOS		MOS	
11B	Infantryman	12B	Combat Engineer
13B	Cannon Crewmember	16S	MANPADS Crewman
19E	M60 Armor Crewman	27E	Tow/Dragon Repairer
19K	M1 Armor Crewman[a]	29E	Comm.-Electronics Radio Repairer
31C	Single Channel Radio Operator	51B	Carpentry/Masonry Specialist
63B	Light-Wheel Vehicle Mechanic	54B	NBC Specialist[d]
71L	Administrative Specialist	55B	Ammunition Specialist
88M	Motor Transport Operator[b]	67N	Utility Helicopter Repairer
91A/B	Medical Specialist/Medical NCO[c]	76Y	Unit Supply Specialist
95B	Military Police	94B	Food Service Specialist
		96B	Intelligence Analyst

[a] Except for the type of tank used, this MOS is equivalent to the 19E MOS originally selected for Project A testing.
[b] This MOS was formerly designated as 64C.
[c] Although 91A was the MOS originally selected for Project A testing, second-tour medical specialists are usually reclassified as 91B.
[d] This MOS was formerly designated as 54E.

Figure 1.2. Project A/Career Force Military Occupational Specialties (MOS)

The MOS in the two groups were carefully sampled to represent the variation in job content in the Army occupational structure. In addition, they were selected so as to overrepresent both the combat specialties and those MOS with the larger proportions of women and minority groups. The MOS selection procedure has been described in detail in previous Project A reports (e.g., Campbell, 1987).

A glossary of terms for the samples and for the different measurement batteries is given in Figure 1.3. The six major samples, their approximate size, and the predictor and/or performance batteries that were to be administered to each are shown in Figure 1.4.

Glossary of Terms	
CVI Sample (CVI)	Soldiers who entered the Army between 1 Jul 83 - 30 Jun 84 <u>and</u> were in the 1985 Project A Concurrent Validation (CVI). They were administered the Trial Predictor Battery and the first-tour job performance measures.
CVII Sample (CVII)	Soldiers who entered the Army between 1 Jul 83 - 30 Jun 84 <u>and</u> were in the 1985 Project A Concurrent Validation (CVI) <u>and</u> the 1988 Second-Tour Concurrent Validation (CVII). They were administered the second-tour job performance measures and were re-administered the ABLE assessment instrument.
LV Sample (LV)	Soldiers in the Longitudinal Validation sample who entered the Army between 20 Aug 86 - 30 Nov 87 <u>and</u> were administered the Experimental Predictor Battery and End-of-Training measures.
LV Training Sample (LVT)	Soldiers in the Longitudinal Validation sample who finished Advanced Individual Training (AIT) and who were administered the End-of-Training measures.
LVI Sample (LVI)	Soldiers who entered the Army between 20 Aug 86 - 30 Nov 87 <u>and</u> were in the LV Sample <u>and</u> the 1988 First-Tour Longitudinal Validation sample (LVI). They were administered the first-tour job performance measures.
LVII Sample (LVII)	Soldiers who entered the Army between 20 Aug 86 - 30 Nov 87 <u>and</u> were in the LV Sample <u>and</u> the LVI Sample <u>and</u> the Longitudinal Validation Second-Tour (LVII) sample. They were administered the second-tour job performance measures in LVII.
Note. Glossary definitions reflect the original research plan. In the course of the research, some CVII soldiers did not have CVI data, some LVI soldiers did not have LV data, and some LVII soldiers did not have both LV and LVI data.	

Figure 1.3. Glossary of Terms for Project A/Career Force Research Samples

Procedure

The data collection procedures for each sample have been described in detail in previous annual reports (e.g., see Campbell & Zook, 1990). Each data collection involved on-site administration by a trained data collection team headed by a team leader from the contractor staff who worked closely with

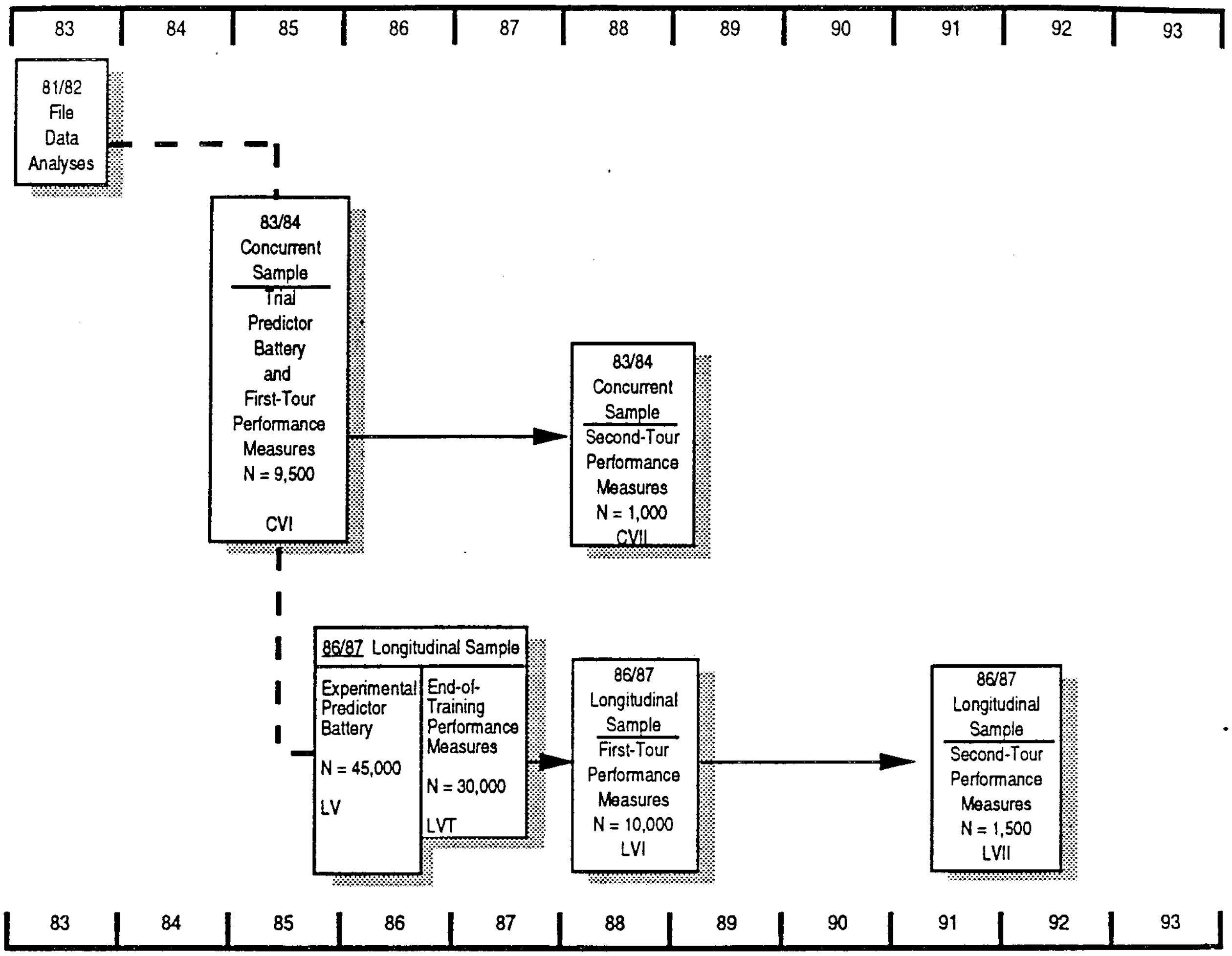


Figure 1.4. Career Force Research Flow and Samples

a designated Army point-of-contact (POC) at the site. A brief characterization of each of the six samples in terms of the timing, location, and duration (per soldier) of the data collection follows.

<u>The Concurrent Validation First-Tour (CVI) sample</u>. The data were collected at 13 posts in the continental United States and at multiple locations in Germany. Each individual was assessed for 1 1/2 days on the project-developed first-tour job performance measures and for 1/2 day on the new predictor measures (the Trial Battery). The individuals in the sample had been in the Army for 12-24 months. Data analysis has been completed for this sample.

<u>The Longitudinal Validation Predictor (LV) Sample</u>. All individuals were assessed on the 4-hour Experimental Predictor Battery within 2 days of first arriving at their assigned Reception Battalion where they would undergo Basic/Advanced Individual training. Data were collected over a 14-month period at eight Reception Battalions by a permanent, on-site data collection team.

The Longitudinal Validation End-of-Training (LVT) Sample. The EOT performance measures were administered to those individuals in the LV sample who completed Advanced Individual Training (AIT), which could take from 2 months to 6 months, depending on the MOS. The training performance measures consisted of an MOS-specific training achievement test and a series of rating scales completed by peers and drill instructors. Data collection took place during the last three days of AIT.

The Longitudinal Validation First-Tour (LVI) Sample. The individuals in the 86/87 cohort who were measured with the Experimental Predictor Battery, completed AIT, and remained in the Army were assessed with the full array of first-tour job performance measures when they were between 12 and 24 months of service. Data collections were conducted at 13 posts in the United States and multiple locations in Europe (primarily in Germany). The administration of the LVI first-tour criterion measures took one day per soldier.

The Concurrent Validation Second-Tour (CVII) Sample. The same data collection teams that administered the first-tour performance measures to the LVI sample also administered the second-tour performance measures at the same location and during the same time periods to a sample of junior NCOs from the 83/84 cohort in their second tour of duty (3-5 years of service). Every attempt was made to include second-tour personnel from the designated MOS who had been part of the first-tour Concurrent Validation sample (CVI). The CVII data collection took one day per soldier.

The Longitudinal Validation Second-Tour (LVII) Sample. The personnel in this sample are members of the 86/87 cohort from the designated MOS who were part of the LV (predictors and training performance measures) and LVI (first-tour job performance measures) samples and who reenlisted for a second tour of duty. The revised second-tour performance measures were administered at 15 U.S. posts, multiple locations in Germany, and two locations in Korea. The LVII performance assessment took one day per soldier.

SUMMARY OF PROJECT EFFORTS FOR YEAR ONE

As described in the first Career Force annual report (Campbell & Zook, 1990), the objectives of the project's first year were focused on developing a full design for the data base and on analyzing basic scores for (a) the final version of the Experimental Predictor Battery (EB), (b) the End-of-Training (EOT) performance measures, and (c) the second-tour criterion measures used to assess NCO performance in the second-tour Concurrent Validation (CVII) sample. The data from the End-of-Training and second-tour Concurrent Validation (CVII) performance assessments were also used to formulate both a model of training performance and a model of second-tour (junior NCO) job performance. That is, the basic scores from the individual performance measures were aggregated into factor scores that represented, as well as possible, the major components, or latent structure, of training performance and second-tour job performance.

By the end of year one, the data collection for the Longitudinal Validation first-tour performance assessments had been completed, but the data cleaning and editing were still in progress and the analysis of the LVI performance measures had not yet begun.

Data Base Design

As described in the first-year annual report, the Career Force data base design allows access at any level of score aggregation. The report describes each variable and the amount of information that is available. The data are accessed via a secure system that requires prior approval by the Army. The data base also includes data from several operational files maintained by the Army.

Basic Scores for the Experimental Battery

During year one, much effort was devoted to analyzing the data that had been obtained by administering the Experimental Predictor Battery to almost 50,000 new accessions in the Longitudinal Validation sample. A number of data editing procedures were compared and evaluated, and great care was taken to maximize data quality for the information that was entered into the final data file. The psychometric properties and subgroup differences for each measure were analyzed, and a series of exploratory and confirmatory analyses were conducted to identify the basic predictor scores within each domain that would be used in the validation analyses.

The final array of tests in the Experimental Battery and the constructs they are intended to measure are shown in Figure 1.5. The 31 basic scores that are obtained from the specific test indicators are shown in Figure 1.6 (Campbell & Zook, 1990).

There was a very high degree of consistency between the Concurrent Validation and the Longitudinal Validation in terms of the factor structures of the various measures. The resulting definitions of the basic predictor scores to be used in the validation analyses were quite similar.

Basic Scores for the End-of-Training Measures

During year one, the data from the school knowledge test and seven training performance rating scales administered at the end of training were analyzed in terms of their psychometric properties and factor structure. Confirmatory techniques were used to identify the "model" of training performance that best represented the covariances among the observed measures. That is, an a priori set of alternative models was proposed and evaluated in terms of the degree to which they fit the data. In the end six basic scores were proposed, two based on the knowledge tests and four based on the rating scales. A brief characterization of the six scores is given in Figure 1.7 (Campbell & Zook, 1990).

These six scores will serve both as criterion measures (for the Experimental Battery) and as predictors (of first-tour and second-tour job performance) in later validation analyses.

Test/Measure	Construct
Paper-and-Pencil Spatial Tests	
Assembling Objects	Spatial Visualization-Rotation
Object Rotation	Spatial Visualization-Rotation
Maze	Spatial Visualization-Scanning
Orientation	Spatial Orientation
Map	Spatial Orientation
Reasoning	Induction
Computer-Administered Tests	
Simple Reaction Time	Reaction Time (Processing Efficiency)
Choice Reaction Time	Reaction Time (Processing Efficiency)
Short-Term Memory	Short-Term Memory
Perceptual Speed and Accuracy	Perceptual Speed and Accuracy
Target Identification	Perceptual Speed and Accuracy
Target Tracking 1	Psychomotor Precision
Target Shoot	Psychomotor Precision
Target Tracking 2	Multilimb Coordination
Number Memory	Number Operations
Cannon Shoot	Movement Judgment
Temperament, Interest, and Job Preference Measures	
Assessment of Background and Life Experiences (ABLE)	Adjustment Dependability Achievement Physical Condition Leadership (Potency) Locus of Control Agreeableness/Likability
Army Vocational Interest Career Examination (AVOICE)	Realistic Interest Conventional Interest Social Interest Investigative Interest Enterprising Interest Artistic Interest
Job Orientation Blank (JOB)	Job Security Serving Others Autonomy Routine Work Ambition/Achievement

Figure 1.5. Experimental Predictor Battery Tests and Relevant Constructs

ASVAB Factor Composites

Quantitative
- Mathematics Knowledge
- Arithmetic Reasoning

Technical
- Auto/Shop Information
- Mechanical Comprehension
- Electronics Information

Speed
- Coding Speed
- Number Operations

Verbal
- Word Knowledge
- Paragraph Comprehension
- General Science

Paper-and-Pencil Test Composite

Spatial
- Assembling Objects Test
- Object Rotation Test
- Maze Test
- Orientation Test
- Map Test
- Reasoning Test

Computer-Administered Test Composites*

Psychomotor
- Target Tracking 1 Distance
- Target Tracking 2 Distance
- Cannon Shoot Time Score
- Target Shoot Distance

Movement Time
- Pooled Movement Time

Perceptual Speed
- Perceptual Speed & Accuracy (DT)
- Target Identification (DT)

Basic Speed
- Simple Reaction Time (DT)
- Choice Reaction Time (DT)

Perceptual Accuracy
- Perceptual Speed & Accuracy (PC)
- Target Identification (PC)

Basic Accuracy
- Simple Reaction Time (PC)
- Choice Reaction Time (PC)

Number Speed and Accuracy
- Number Memory (Operation DT)
- Number Memory (PC)

Short-Term Memory
- Short-Term Memory (PC)
- Short-Term Memory (DT)

ABLE Composites

Achievement Orientation
- Self-Esteem
- Work Orientation
- Energy Level

Leadership Potential
- Dominance

Dependability
- Traditional Values
- Conscientiousness
- Nondelinquency

Adjustment
- Emotional Stability

Cooperativeness
- Cooperativeness

Internal Control
- Internal Control

Physical Condition
- Physical Condition

JOB Composites

High Job Expectations
- Pride
- Job Security
- Serving Others
- Ambition

Job Routine
- Routine

Job Autonomy
- Autonomy

AVOICE Composites

Rugged/Outdoors
- Combat
- Rugged Individualism
- Firearms Enthusiast

Audiovisual Arts
- Drafting
- Audiographics
- Aesthetics

Interpersonal
- Medical Services
- Leadership/Guidance

Skilled/Technical
- Science/Chemical
- Computers
- Mathematics
- Electronic Communication

Administrative
- Clerical/Administrative
- Warehousing/Shipping

Food Service
- Food Service - Professional
- Food Service - Employee

Protective Services
- Fire Protection
- Law Enforcement

Structural/Machines
- Mechanics
- Heavy Construction
- Electronics
- Vehicle Operator

*DT = Decision Time and PC = Proportion Correct

Figure 1.6. Longitudinal Validation Experimental Battery: Composite Scores and Constituent Basic Scores

EOT RATING SCALE BASED SCORES

1) **Effort and Technical Skill (ETS)**

Technical Knowledge/Skill:	How effective is each soldier in acquiring job/soldiering knowledge and skill?
Effort:	How effective is each soldier in displaying extra effort?

2) **Maintaining Personal Discipline (MPD)**

Following Regulations and Orders:	How effective is each soldier in adhering to regulations, orders, and SOP and displaying respect for superiors?
Self Control:	How effective is each soldier in controlling own behavior related to aggressive acts?

3) **Physical Fitness and Military Bearing (PFB)**

Military Appearance:	How effective is each soldier in maintaining proper military appearance?
Physical Fitness:	How effective is each soldier in maintaining military standards of physical fitness?

4) **Leadership Potential (LEAD):**

Leadership Potential:	Evaluate each soldier on his or her potential effectiveness as a leader. Do not necessarily rate on the basis of present performance.

EOT KNOWLEDGE TEST BASED SCORES

5) **Basic Knowledge Score:**	Items measuring knowledge requirements common to all MOS.
6) **Technical Knowledge Score:**	Items measuring technical knowledge requirements specific to each MOS.

Figure 1.7. Composite Scores That Reflect End-of-Training (EOT) Performance Factors

Development of Second-Tour Performance Scores (CVII)

The performance measures used in the CVII sample, and their development, have been described in detail in previous Project A and Career Force reports (e.g., Campbell, 1991; Campbell & Zook, 1991). First-tour measures were revised for use with second-tour personnel and new measures reflecting the unique components of second-tour jobs were added. A summary description of the specific measures is given below.

Rating Scales

On the basis of second-tour critical incident analyses, the Army-wide Behaviorally Anchored Ratings Scales (BARS) and MOS-specific BARS were revised and scales having to do with leadership and supervision were added. Further, based on job analysis data, seven new scales pertaining to supervision and leadership responsibilities were also added. A full list of the Army-wide rating scales is shown below. Not shown are the MOS BARS for each MOS, which were revised to reflect second-tour performance demands, and the Combat Performance Prediction Scales, which were the same as those used in LVI, and which were not administered to female NCOs.

Army-Wide Behavior Scales:

1. Demonstrating Technical Knowledge and Skill
2. Demonstrating Effort
3. Supervising Subordinates
4. Following Regulations and Orders
5. Demonstrating Integrity
6. Training and Development of Subordinates
7. Maintaining Equipment
8. Physical Fitness
9. Self-Development
10. Showing Consideration for Subordinates
11. Demonstrating Appropriate Military Bearing
12. Demonstrating Appropriate Self-Control

Additional Leadership Scales:

13. Serving as a Role Model
14. Communication With Subordinates
15. Personal Counseling
16. Monitoring Subordinate Performance
17. Organizing Missions/Operations
18. Personnel Administration
19. Performance Counseling

General Scales:

20. Overall Effectiveness
21. Senior NCO Potential

Situational Judgment Test (SJT)

A new paper-and-pencil measure of supervisory judgment was developed by describing prototypical judgment situations and asking the respondent to select the most appropriate and the least appropriate courses of action. The situation descriptions and the scoring keys were refined through extensive subject matter expert (SME) judgments.

Supervisory Simulation Exercises

These measures were developed to assess NCO performance in job areas that were judged to be best assessed through the use of interactive exercises. The simulations were designed to evaluate performance in counseling and training subordinates. A trained evaluator (role player) played the part of a subordinate to be counseled or trained and the examinee assumed the role of a first-line supervisor who was to conduct the counseling or training. Evaluators also scored the examinee's performance, using a standard set of rating scales.

Here are brief descriptions of the three simulation exercises:

- **Personal Counseling Simulation**: A private first class (PFC) is exhibiting declining job performance and personal appearance. Recently, the PFC's wall locker was left unsecured. The supervisor has decided to counsel the PFC about these matters.

- **Disciplinary Counseling Simulation**: There is convincing evidence that the PFC lied to get out of coming to work today. The PFC has arrived late to work on several occasions and has been counseled for lying in the past. The PFC has been instructed to report to the supervisor's office immediately.

- **Training Simulation**: The commander will be observing the unit practice formation in 30 minutes. The private, although highly motivated, is experiencing problems with the hand salute and about face.

For each exercise, examinee performance was evaluated on 3-point rating scales reflecting specific behaviors tapped by the exercises and a 5-point overall effectiveness rating scale. Factor analyses of the ratings data suggested that each simulation could be scored in terms of the content of the NCO's behavior (i.e., did he or she do or say the right things) and the process, or style, with which the counseling steps were carried out.

Administrative Measures

The self-report Personnel File Form (PFF) used in LVI was modified for use with second tour and six administrative indices of performance were obtained.

Job Knowledge and Hands-On Measures

The content of each of these measures was revised on the basis of the second-tour job analyses and the revised instruments were subjected to extensive SME review. Analyses of alternative aggregations of item and scale scores from both of these measures resulted in the adoption of a general (Army-wide) and an MOS-specific score for each of them.

Final Array of Second-Tour Basic Performance Scores

After extensive analyses of their psychometric properties and factor structures, based on CVII data, the final array of basic second-tour performance scores was as shown in Figure 1.8. There were 22 basic scores. Scores from this array became the basis for the second-tour performance modeling analysis in CVII.

Development of the CVII Second-Tour Performance Model

The basic CVII performance scores served as input to the development of a latent structure model for second-tour performance (Campbell & Zook, 1990). Based on a consensus of the project staff, three major alternatives could be used to explain the observed correlations. Consequently, the competing models that were evaluated for comparative goodness of fit, using the LISREL VI program (Jöreskog & Sörbom, 1986), were the following:

(1) First-Tour Model: Included five substantive and two methods factors, with the SJT and Simulation variables all loading on the Effort and Leadership factor.

(2) Leadership Factor Model: Included a sixth substantive factor with the SJT, Simulation, and Leadership Rating factor variables all loading on this factor. This model was evaluated with and without a separate simulation "methods" factor.

(3) Training and Counseling Factor Model: Included a sixth substantive factor with just the Simulation variables. No separate simulation methods factor could be estimated under this model.

Of the three models, the Training and Counseling Factor Model provided the closest fit to the observed data. A result of considerable interest was that the SJT (a paper-and-pencil measure) fit best with the Effort and Leadership factor, in spite of the method variance involved.

The basic scores that have been used to represent the latent variables are as shown in Figure 1.9. For validation analysis purposes, the six substantive factor scores are obtained by standardizing and summing the basic scores within each factor.

Hands-On Performance Test

1. MOS-specific task performance score
2. General (common) task performance score

Job Knowledge Test

3. MOS-specific task knowledge score
4. General (common) task knowledge score

Army-Wide Rating Scales

5. Leadership/supervision composite
6. Technical skill and effort composite
7. Personal discipline composite
8. Physical fitness and military bearing composite

MOS-Specific Rating Scales

9. Overall MOS composite

Combat Performance Prediction Scales

10. Overall Combat Prediction scale composite (available for males only)

Personnel File Form

11. Awards and Certificates
12. Articles 15/Flag Actions (Disciplinary Actions)
13. Physical Readiness
14. M16/M19 Qualification
15. Military Training Courses
16. Promotion Rate

Situational Judgment Test

17. Total score obtained by subtracting the total "ineffectiveness" score from the total "effectiveness" score

Supervisory Simulation Exercises

18. Personal Counseling: Process
19. Personal Counseling: Content
20. Disciplary Counseling: Process
21. Disciplary Counseling: Content
22. Training: Total composite score

Figure 1.8. Summary List of CVII Basic Criterion Scores

Latent Variables in the CVII Performance Model

- **Core Technical Proficiency (CTP)**
  - MOS-Specific Hands-On
  - MOS-Specific Job Knowledge

- **General Soldiering Proficiency (GSP)**
  - General (Common) Hands-On
  - General (Common) Job Knowledge

- **Effort and Leadership (ELS)**
  - Awards and Certificates
  - Military Training Courses
  - Promotion Rate
  - Leadership/Supervision Rating Composite
  - Technical Skill/Effort Rating Composite
  - Overall MOS Rating Composite
  - Situational Judgment Test Total Score

- **Personal Discipline (MPD)**
  - Disciplinary Actions (reversed)
  - Personal Discipline Rating Composite

- **Physical Fitness/Military Bearing (PFB)**
  - Physical Readiness Score
  - Physical Fitness/Bearing Rating Composite

- **Training and Counseling Subordinates (TCS)**
  - Simulation Exercise - Personal Counseling Content
  - Simulation Exercise - Personal Counseling Process
  - Simulation Exercise - Disciplinary Content
  - Simulation Exercise - Disciplinary Process
  - Simulation Exercise - Training

- **Written Methods (WM)**
  - MOS-Specific Knowledge
  - Common Soldiering Knowledge
  - Situational Judgment Test

- **Ratings Methods (RM)**
  - Four Army-Wide Rating Composites
  - Overall MOS Rating Composite

Figure 1.9. Relationship of Specific Variables to Overall Factors in the CVII Performance Model.

SUMMARY OF PROJECT EFFORTS FOR YEAR TWO

As described in the annual report for the 1991 fiscal year (Campbell & Zook, 1994a), year two was a period of score development, model building, and basic validation analyses for (a) training performance (EOT), (b) first-tour performance (LVI), and (c) second-tour performance (CVII). During year two, the second-tour longitudinal data collection (LVII) began and was ongoing.

Objectives

The specific objectives for the second-year annual report were as follows:

(1) Describe the development of alternative scores for the Assessment of Background and Life Experiences (ABLE) instrument.

(2) Describe the basic validation analyses for the prediction of performance in training.

(3) Describe the development of basic scores for the longitudinal sample first-tour performance measures.

(4) Describe the replication/confirmation of the first-tour performance model and the basic Longitudinal Validation analyses for the Experimental Predictor Battery against first-tour performance.

(5) Describe the basic validation analyses for the prediction of second-tour performance, using the CVII sample.

(6) Report the results of a preliminary analysis of the prediction of second-tour performance from first-tour predictors and performance.

Development of Alternative ABLE Factor Composites

As part of Project A, and based on the results of an extensive review of the literature, 10 temperament scales had been developed to form the ABLE. These constructs were selected as the most promising for predicting performance in Army enlisted occupational specialties. In addition, four validity scales were included to detect inaccuracies in self-reports of temperament and a self-report measure of physical condition was also included (see Hough, Eaton, Dunnette, Kamp, & McCloy, 1990, for more information on the development of ABLE). To develop a set of conceptually meaningful construct (composite) scores, Peterson et al. (1990) carried out both exploratory and confirmatory factor analyses on the correlation among the content scale scores.

The resulting seven temperament constructs (composites) and associated ABLE scales are shown in Table 1.1. The constructs of Dependability, Dominance (Surgency), Adjustment, and Cooperativeness have counterparts in the Big Five personality dimensions described by Norman (1963) and Goldberg (1981). Conversely, Achievement and Internal Control are not in the Big Five taxonomy, but were among the strongest predictors of job performance in the Project A review of the temperament domain (see Hough, 1992, for more details on the relationship of ABLE constructs to the Big Five).

Table 1.1

ABLE Rational Composites and Corresponding Content Scales

Composite	ABLE Scale
Achievement Orientation	Self-Esteem Work Orientation Energy Level
Leadership Potential	Dominance
Dependability	Traditional Values Conscientiousness Nondelinquency
Adjustment	Emotional Stability
Cooperativeness	Cooperativeness
Internal Control	Internal Control
Physical Condition	Physical Condition

As noted above, a rational/theoretical approach was the primary method used in developing ABLE. An alternative empirical procedure emphasizes the internal covariance structure of a set of items and uses factor analytic methods. Consequently, during year two, internal scale construction methods were used to increase, through homogeneous keying, the internal consistency of ABLE composites and to decrease their intercorrelations.

Results from factor analyses of 199 items were used to form seven preliminary composites. These composites contained 99 items. Next, correlations between the remaining content-type items (excluding the validity scale items) and the preliminary factor composites were examined and each remaining item was assigned to the composite with which it had the highest correlation. The seven factor composites resulting from this procedure used 168 items and are called the ABLE-168 composites. In all, 125 items were assigned in the same way on the ABLE-168 composites and the ABLE rational composites.

As a second alternative, an item was retained only if it correlated at least .33 with the scale for which it was assigned and had a higher correlation with its own composite (by .03) than any other. In addition, several items that added only minimally to internal consistency were dropped. The resulting set of composites had a total of 114 items and is called the ABLE-114 composites. Eighty-nine of these items were assigned in the same way on ABLE-114 and the ABLE rational composites.

The three scoring methods converged to yield seven similar temperament constructs. The composites measuring the same constructs were very highly correlated ($\underline{r}$ = .88 to 1.0).

ABLE-114 composites had greater discriminant validity than either the ABLE-168 factor composites or the ABLE rational composites. The average correlation among the composites (off-diagonal elements) was .40 for ABLE-114, and .47 for the ABLE rational composites and ABLE-168.

Table 1.2 shows the distribution of items on ABLE-168 and ABLE-114 for each of the ABLE content scales. Items outside the shaded areas were assigned differently on the rational and factor composites. There is much overlap between the rational and factor composites. However, approximately 25 percent of item assignments for the factor composites were different from those used for the rational composites. Most of these are consistent with results from previous research and/or can be understood on the basis of item content.

In sum, there are three alternative sets of ABLE composites measuring seven temperament constructs. The 114-item form is shorter and has higher discriminant validity than the other two sets of composites, with little apparent loss of reliability. Subsequent analyses in the Career Force Project examine the criterion-related validities of these alternative sets of composites.

Prediction of Performance in Training

The objectives of analyses of the end-of-training (EOT) data were to:

(1) Compute the validities for ASVAB and Experimental Battery predictors against rating measures and also paper-and-pencil test measures of training performance.

(2) Compare the validities of four alternative sets of ASVAB scores.

(3) Compare the validities of three alternative sets of ABLE scores.

(4) Assess the incremental validities for the Experimental Battery predictors over ASVAB.

Procedure

The EOT validation analysis consisted of the following steps:

A) Multiple correlations between each set of predictor scores and each set of criterion scores were computed separately by MOS and then averaged across the Batch A MOS and across all MOS.

Table 1.2

Distribution of ABLE Scale Items on ABLE-168 and ABLE-114 Factor Composites

ABLE Scale	No. of Items	ABLE Factor Composite						
		Achievement Orientation	Leadership Potential	Dependability	Adjustment	Cooperativeness	Internal Control	Physical Condition
Self-Esteem	12 (6)		10 (6)		2 (0)			
Work Orientation	19 (15)	18 (14)	1 (1)					
Energy Level	21 (9)	13 (6)			6 (1)			2 (2)
Dominance	12 (12)		12 (12)					
Traditional Values	11 (7)			5 (4)			5 (3)	
Conscientiousness	15 (11)	9 (8)		6 (3)				
Nondelinquency	20 (13)			20 (13)				
Emotional Stability	17 (11)				17 (11)			
Cooperativeness	18 (11)			2 (1)		16 (10)		
Internal Control	16 (11)	2 (0)			2 (1)		12 (10)	
Physical Condition	6 (6)							6 (6)
Poor Impression	2 (2)				2 (2)			
Total	169 (114)	42 (28)	23 (19)	33 (21)	29 (15)	16 (10)	17 (13)	8 (8)

Note. ABLE-114 items are shown in parentheses. Shaded areas indicate convergence between the rational and factor composites.

1) The ASVAB predictor set was represented by:

   a) The nine ASVAB subtest scores
   b) The four ASVAB factor scores
   c) The Armed Forces Qualification Test (AFQT)
   d) The MOS-appropriate Aptitude Area composite score

2) The ABLE predictor set was represented by three sets of scores:

   a) The seven rational scales
   b) Seven empirical scales that retained 168 items
   c) Seven empirical scales that retained only 114 items

3) Each of the other predictor sets (i.e., spatial, computer, AVOICE, JOB) was represented as in previous analyses.

All results were adjusted for shrinkage and corrected for multivariate range restriction.

B) Incremental validity was computed for each set of Experimental Battery predictors over the ASVAB.

C) Multiple correlations were computed between each set of predictor scores and a "Peer 1" rating, a "Peer 2" rating, a supervisor rating, and various combinations.

Results

To summarize the principal findings, multiple correlations for six predictor sets are shown in Table 1.3; the incremental validities are summarized in Table 1.4. In general, ASVAB shows high validity against the school knowledge measures and the relative validities for the four ratings factors are as would be expected on the basis of the factors. The ABLE does not predict the "will do" factors quite as well as it did in CVI but it predicts the "can do" factors somewhat better.

These results indicate that the level of validity of the ASVAB factors for predicting the School Knowledge (SK) test scores was extremely high, especially for the Technical (SK-Tech) and Total (SK-Total) scores. Likewise, the spatial composite and the computer battery produced high validities for these criteria.

Results from other analyses indicate that peer ratings of training performance are more accurately predicted than supervisor ratings of training performance. This suggests that peer ratings may be more valid training measures than supervisor ratings, presumably because, in training, peers generally have greater opportunity to observe ratees than do supervisors. This comparison is confounded, however, by the greater reliability of the peer ratings that is, at least in part, due to the fact that they are based on more raters per ratee than are the supervisor ratings. Yet analyses at the 1-rater level corroborate the notion that the peer ratings have more utility than the supervisor ratings for assessing training performance.

Table 1.3

Mean of Multiple Correlations Computed Within Job for End-of-Training Sample for ASVAB Factors, Spatial, Computer, JOB, ABLE Rational Composites, and AVOICE

Criterion[a]	MOS	No. of MOS[b]	ASVAB Factors [4]	Spatial [1]	Computer [8]	JOB [3]	ABLE Comp. [7]	AVOICE [8]
Peer-ETS	Batch A	11	41 (07)	35 (05)	36 (05)	24 (06)	19 (09)	22 (07)
	All MOS	22	43 (13)	37 (10)	33 (14)	23 (11)	23 (12)	23 (10)
Peer-MPD	Batch A	11	25 (04)	22 (05)	21 (05)	09 (07)	19 (05)	11 (07)
	All MOS	22	26 (11)	22 (08)	15 (10)	12 (10)	22 (10)	09 (09)
Peer-PFB	Batch A	11	14 (09)	05 (06)	11 (05)	05 (05)	29 (06)	07 (07)
	All MOS	22	19 (14)	10 (11)	12 (09)	09 (12)	26 (11)	10 (10)
Peer-LEAD	Batch A	11	30 (10)	24 (07)	28 (07)	18 (09)	22 (09)	17 (10)
	All MOS	22	30 (16)	26 (12)	25 (16)	20 (14)	22 (12)	16 (14)
Supv-ETS	Batch A	11	21 (06)	18 (05)	17 (10)	10 (08)	09 (10)	11 (10)
	All MOS	22	27 (15)	22 (11)	18 (13)	10 (10)	11 (12)	10 (10)
Supv-MPD	Batch A	11	13 (09)	12 (07)	11 (08)	06 (06)	05 (06)	06 (06)
	All MOS	22	16 (16)	14 (11)	10 (13)	06 (08)	05 (07)	04 (06)
Supv-PFB	Batch A	11	11 (07)	09 (05)	09 (08)	06 (05)	11 (09)	07 (07)
	All MOS	22	16 (15)	13 (12)	11 (15)	05 (07)	11 (11)	05 (06)
Supv-LEAD	Batch A	11	15 (10)	14 (08)	13 (10)	08 (08)	10 (11)	08 (09)
	All MOS	22	19 (17)	17 (11)	12 (12)	11 (09)	11 (12)	07 (09)
SK-Basic	Batch A	9	68 (06)	57 (06)	57 (06)	38 (05)	30 (07)	37 (05)
	All MOS	20	67 (08)	58 (07)	55 (14)	36 (10)	31 (14)	37 (11)
SK-Tech	Batch A	11	76 (05)	63 (05)	61 (05)	41 (07)	33 (05)	44 (07)
	All MOS	22	75 (06)	62 (08)	59 (06)	38 (11)	33 (13)	40 (12)
SK-Total	Batch A	11	78 (03)	65 (04)	64 (03)	43 (07)	34 (05)	45 (06)
	All MOS	22	77 (05)	65 (07)	62 (07)	40 (11)	35 (14)	42 (13)

Note: Corrected for range restriction and adjusted for shrinkage (Rozeboom, 1978, formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] ETS = Effort and Technical Skill; MPD = Maintaining Personal Discipline; PFB = Physical Fitness and Military Bearing; LEAD = Leadership Potential; SK = School Knowledge.
[b] Number of MOS for which validities were computed.

Table 1.4

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for End-of-Training Sample for Spatial, Computer, JOB, ABLE Rational Composites, and AVOICE

Criterion[a]	MOS	No. of MOS[b]	A4 ASVAB Factors [4]	A4+ Spatial [5]	A4+ Computer [12]	A4+ JOB [7]	A4+ ABLE Comp. [11]	A4+ AVOICE [12]
Peer-ETS	Batch A	11	*41* (07)	42 (07)	42 (06)	41 (07)	44 (06)	41 (07)
	All MOS	22	*43* (13)	42 (14)	40 (16)	42 (13)	45 (11)	41 (14)
Peer-MPD	Batch A	11	*25* (04)	25 (05)	24 (05)	25 (05)	34 (06)	24 (07)
	All MOS	22	*26* (11)	25 (11)	22 (12)	25 (12)	33 (11)	22 (11)
Peer-PFB	Batch A	11	*14* (09)	13 (09)	17 (07)	15 (09)	31 (09)	15 (09)
	All MOS	22	*19* (14)	18 (14)	16 (12)	19 (17)	30 (14)	18 (11)
Peer-LEAD	Batch A	11	*30* (10)	30 (10)	31 (08)	30 (11)	35 (09)	29 (13)
	All MOS	22	*30* (16)	30 (17)	28 (18)	31 (18)	34 (15)	28 (18)
Supv-ETS	Batch A	11	*21* (06)	21 (07)	19 (09)	20 (06)	19 (12)	17 (12)
	All MOS	22	*27* (15)	26 (15)	24 (15)	25 (15)	25 (19)	22 (16)
Supv-MPD	Batch A	11	*13* (09)	12 (09)	11 (09)	11 (09)	13 (11)	11 (10)
	All MOS	22	*16* (16)	16 (16)	12 (17)	14 (17)	16 (16)	11 (14)
Supv-PFB	Batch A	11	*11* (07)	11 (07)	10 (08)	10 (07)	16 (09)	10 (09)
	All MOS	22	*16* (15)	15 (14)	12 (15)	14 (13)	18 (13)	11 (13)
Supv-LEAD	Batch A	11	*15* (10)	14 (10)	14 (11)	14 (10)	16 (13)	13 (12)
	All MOS	22	*19* (17)	19 (17)	15 (15)	19 (16)	20 (17)	15 (15)
SK-Basic	Batch A	9	*68* (06)	69 (06)	68 (06)	68 (06)	68 (07)	68 (06)
	All MOS	20	*67* (08)	68 (08)	65 (16)	67 (09)	66 (11)	66 (10)
SK-Tech	Batch A	11	*76* (05)	77 (05)	77 (05)	76 (05)	76 (05)	76 (05)
	All MOS	22	*75* (06)	75 (06)	75 (05)	75 (06)	75 (07)	74 (07)
SK-Total	Batch A	11	*78* (03)	79 (03)	79 (03)	78 (03)	79 (03)	78 (04)
	All MOS	22	*77* (05)	77 (05)	77 (05)	77 (05)	77 (06)	76 (06)

Note: Corrected for range restriction and adjusted for shrinkage (Rozeboom, 1978, formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Multiple $\underline{R}$s for ASVAB Factors alone are in italics. Underlined numbers denote multiple $\underline{R}$s greater than for ASVAB Factors alone. Decimals omitted.

[a] ETS = Effort and Technical Skill; MPD = Maintaining Personal Discipline; PFB = Physical Fitness and Military Bearing; LEAD = Leadership Potential; SK = School Knowledge.
[b] Number of MOS for which validities were computed.

Further analysis showed that the average multiple correlations for the four different sets of ASVAB scores differed only slightly in validity, except that the peer ratings of Physical Fitness (PFB) were better predicted by the nine subtests and the four factors. However, the school knowledge test scores were predicted somewhat better (about three to five points) by the ASVAB subtests and factors than by the AFQT or Aptitude Area composites.

Both ABLE and AVOICE predicted the knowledge-based scores quite well. The largest incremental validities were for ABLE over ASVAB when predicting Personal Discipline, Fitness and Bearing, and Leadership.

Finally, there were virtually no differences in validities for the three alternative sets of ABLE scores although the ABLE-114 validities were consistently slightly higher.

Development of Basic Scores for the Longitudinal Validation (LVI) Performance Measures

In 1988 and 1989, first-tour criterion measures were administered to the Longitudinal Validation sample (LVI). This data collection was conducted concurrently with the administration of second-tour criterion measures to the Concurrent Validation sample (CVII). Before the LVI performance model development and subsequent validation analyses could begin, it was necessary to derive basic scores for each of the individual first-tour job performance measures. Dealing with all the individual scores from each task test, each rating scale, and each administrative index was simply not feasible or desirable. There were too many, and the reliabilities of the individual items or scales preserved too much measurement error with very little gain in total information. Consequently, the full array of scale scores was aggregated into a smaller set of basic scores for each measure.

Table 1.5 lists the individual measures that were administered.

Differences Between CVI and LVI Performance Measures

The 3-year time period between CVI and LVI raised the issue that for the job knowledge and hands-on measures, equipment and/or procedural changes would require test revisions, and changes in MOS responsibilities had the potential of making some tasks obsolete.

Project staff identified relevant changes so that the appropriate revisions could be made. In a few cases where an entire task was obsolete, the task was dropped without replacement. In many cases, revisions were simply a matter of replacing outdated terminology. Updated criterion measures were forwarded to the MOS proponents for a currency review and additional revisions were made on the basis of this review.

While there was considerable interest in keeping the Combat Performance Prediction Scales, project staff and the Scientific Advisory Group agreed that the version used in CVI was too lengthy. New scales were field tested in conjunction with the second-tour criterion measure field tests. The decision was made to retain the original summated scale format, but the total number of items was reduced from 40 to 19.

Table 1.5

Measures Administered to Soldiers in LVI Sample

MOS in Batch A:	Background Information Form Job Knowledge Tests Hands-On Tests Army-Wide Rating Scales MOS-Specific Rating Scales Combat Performance Prediction Scales (males only) Personnel File Form Army Job Satisfaction Questionnaire Job History Questionnaire Physical Requirements Survey
MOS in Batch Z:	Background Information Form School Knowledge Test Army-Wide Rating Scales Combat Performance Prediction Scales (males only) Personnel File Form Army Job Satisfaction Questionnaire Physical Requirements Survey

Note. Rating scale data were collected from both supervisors and peers. The Physical Requirements Survey is not a Career Force or Project A measure.

The self-report form for gathering information on administrative records was updated by reviewing its contents with officers and NCOs representing the Army Personnel Command (PERSCOM). The form was altered to allow soldiers to report an M19 qualification in the event that an M16 qualification was not applicable. Also, three awards were dropped per guidance from PERSCOM.

Task-level ratings were deleted from the array of Batch A first-tour criterion measures used in CVI. The Army-wide and MOS-specific rating scales were retained in their original form.

The development of the basic scores for each measure was based on the performance data collected from individuals in the Batch A and Batch Z MOS that were included in the administration of first-tour criterion measures in 1988 and 1989. The Batch A MOS were the same as those studied in the Concurrent Validation, except for the addition of 19K (M1 Armor Crewman).

As in CVI, the Batch A MOS differed from the Batch Z MOS in the comprehensiveness of the MOS-specific criterion measures that were available for administration. MOS-specific rating scales, hands-on tests, and job knowledge tests were administered to Batch A soldiers. The only MOS-specific measure available for administration to the Batch Z soldiers was the school

knowledge test that had been developed for administration at the end of training. The school knowledge test was administered to the Batch Z examinees as a surrogate for a job knowledge test.

Score Development for Administrative Indices

Five scores were computed from the LVI Personnel File Form: (a) awards and memoranda/certificates of achievement, (b) Physical Readiness Test, (c) M16 qualification, (d) Articles 15 and flag actions (disciplinary actions), and (e) promotion rate.

The first score was a composite of (a) awards and decorations; (b) memoranda of appreciation, commendation, or achievement; and (c) certificates of appreciation, commendation, or achievement. The last score, promotion rate, was derived from data available in the Army's computerized personnel records. It was the residual of pay grade regressed on time in service, adjusted by MOS.

Basic Score for the Combat Performance Prediction Ratings

Principal components analyses of the LVI/CVII Combat Scale data indicated the presence of two factors. The second factor, however, was defined by the three negatively worded items. Given that the second factor was probably not substantively distinct from the first, a single total score (with the negatively worded items reverse-scored) was calculated for the Combat Scale ratings. (The two factors were essentially the same as those found in CVI, where two Combat Scale scores were derived.)

Development of Basic Scores for the First-Tour Performance Rating Scales

The Army-wide rating scales include 12 dimensions of soldier effectiveness that are important regardless of soldiers' MOS. MOS-specific rating scales were developed for each of the nine Batch A MOS, and these rating scales include between 7 and 13 dimensions of MOS-specific performance.

Principal factor analyses with varimax rotation were conducted on the Army-wide ratings (across all MOS), for supervisor and peer ratings separately *and* pooled together. The pooled ratings were computed by averaging the mean peer rating and one supervisor rating for those soldiers who had at least one peer rating and one supervisor rating. Because previous analyses (using the CVI sample) showed that a single factor was sufficient to account for the majority of the variance in the MOS-specific ratings, factor analyses were not conducted for the MOS-specific rating data.

Table 1.6 shows the three-factor, rotated solutions for the pooled peer/supervisor ratings. These data demonstrate the remarkable similarity of the rotated factor structures for the CVI and LVI samples. It is worth noting that these same three factors were also obtained in factor analyses of performance rating data for a sample of 950 second-tour soldiers, which was collected using a set of rating scales very similar to those used to collect the present data (Campbell & Zook, 1990).

Table 1.6

Comparison of LVI and CVI Army-Wide Factor Analysis[a] Results: Pooled Peer/Supervisor Ratings[b]

	Factor Loadings (LVI/CVI)		
Dimension	1	2	3
Technical Knowledge/Skill	.67/.71	.30/.28	.38/.30
Leadership	.65/.69	.34/.30	.44/.37
Effort	.66/.69	.47/.43	.32/.26
Self-Development	.52/.57	.42/.38	.46/.38
Maintaining Equipment	.50/.54	.41/.34	.41/.35
Following Regulations	.39/.41	.73/.69	.31/.30
Self-Control	.19/.22	.65/.63	.20/.20
Integrity	.44/.50	.66/.59	.30/.28
Military Bearing	.31/.32	.35/.32	.57/.57
Physical Fitness	.24/.21	.16/.15	.49/.49
Percent Common Variance	37.7/44.9	36.6/32.7	25.6/22.4

Note. Sample size is 7,919 for LVI and 8,642 for CVI.

[a] Principal factor analysis, varimax rotation.
[b] Computed by averaging the mean peer rating and the mean supervisor rating.

For both the Army-wide and MOS-specific rating scales, the mean, variability, and reliability of the peer, supervisor, and pooled peer/supervisor ratings appear quite acceptable and are comparable to what was found in the CVI research. Factor analyses of the Army-wide ratings showed that the three-factor CVI solution was replicated in the present data. Accordingly, the three composites shown in Table 1.7, along with the overall effectiveness rating, were used as the basic scores for the Army-wide rating data.

Table 1.7

Composition and Definition of LVI Army-Wide Rating Composites

Factor Name and Definition	Percent Common Variance Accounted For by Relevant Factor[a] (LVI/CVI)	Dimensions Included
1. Technical Skills and Job Effort: Exerting effort over the full range of job tasks; engaging in training or other development activities to increase proficiency; persevering under dangerous or adverse conditions; and demonstrating leadership and support toward peers.	37.8/44.9	Technical Knowledge/ Skill Leadership Effort Self-Development Maintaining Equipment
2. Personal Discipline: Adhering to Army rules and regulations; exercising self-control; demonstrating integrity in day-to-day behavior; and not causing disciplinary problems.	36.6/32.7	Following Regulations Self-Control Integrity
3. Physical Fitness/Military Bearing: Maintaining an appropriate military appearance and bearing and staying in good physical condition.	25.6/22.4	Military Bearing Physical Fitness

[a] Factor analysis of pooled peer/supervisor ratings.

Development of Basic Scores for Hands-On Performance and Job Knowledge Measures

As the first step in replicating the CVI procedures for constructing the basic scores, tasks were clustered into Functional Categories as described in the Project A annual report for 1986 (Campbell, 1987).

Following the procedures developed with the CVI data, tasks were also sorted into six higher level groups referred to as Task Factors (Communication, Vehicles, Basic Techniques, Identify Targets, Technical, and Safety/Survival) and known as CVBITS. Tasks were also combined into just two groups: General (i.e., Army-wide) and MOS-specific.

In general, the grouping schemes are hierarchical: Tasks (the lowest level) are placed in Functional Categories, the Functional Categories (level two) are aggregated to form the six Task Factors (level three), and Task Factors are then aggregated to form the two Task Constructs (level four), as diagrammed in Figure 1.10.

For the LVI data, confirmatory factor analyses were conducted to assess the fit of alternative levels of score aggregation. These analyses served two purposes: They were used to assess the relative merits of each model and to corroborate the CVI decision to use the six task factor scores (CVBITS). The analysis required the computation of separate tests of goodness of fit for hands-on and job knowledge test data, for each of the 10 MOS, on each of three competing models. The three models tested were: a one-factor model, postulating the existence of a single factor in the data; a two-factor model, proposing the Basic and the Technical Task Constructs; and a three-to-six-factor model (the number of factors varying among MOS and test method), using the Task Factors. Examination of the results from LVI argues for the retention of the six Task Factor scores for both the Hands-On and Job Knowledge measures.

Final Array of LVI Basic Performance Scores

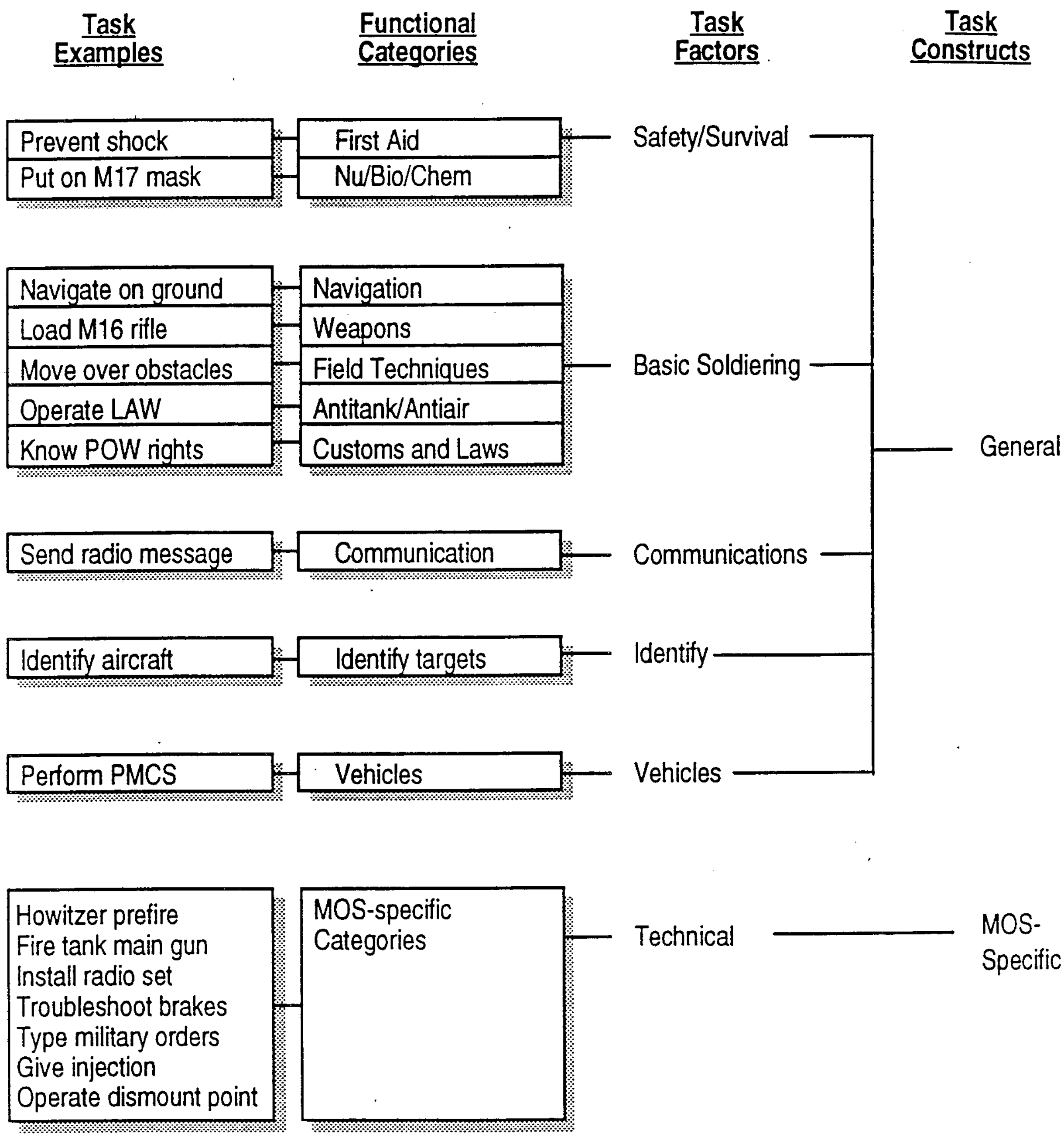
A summary list of the basic performance scores produced by the analyses summarized above is given in Figure 1.11. These are the scores that were put through the final editing and score imputation procedures for the LVI data file. The scores that formed the basis for the confirmatory tests of the LVI model of first-tour job performance were also drawn from this array.

The LVI Data File: Final Data Editing and Score Imputation

The Longitudinal Validation First-Tour (LVI) data were collected from 11,266 soldiers in 21 MOS -- 6,815 from Batch A MOS and 4,451 from Batch Z MOS. Extensive efforts were made to collect complete information from each examinee for all instruments. However, as with all data collection exercises, circumstances precluded complete success. The final counts of soldiers for whom data were analyzed for each instrument are given in Tables 1.8 and 1.9 for Batch A and Batch Z MOS, respectively.

Data for each performance measure were processed individually. After processing was completed for these individual measures, they were combined so that all LVI data for each examinee were included in a single file. The data were combined separately by MOS. When the data were combined, basic scores were calculated for the individual performance measures. Table 1.10 shows the amount of missing data for the final set of basic criterion scores.

In addition to the performance data, missing Longitudinal Validation predictor data were also imputed. For a complete description of the editing process used on the predictor data, see the 1990 annual report. The bulk of the editing process was accomplished during FY90, but additional work was done during FY91. The amounts of missing data for each score on each paper-and-pencil and each computerized measure are shown in Tables 1.11 and 1.12.



Note. The Task Factors correspond to the six task groups known as CVBITS. The Task Constructs termed General and MOS-Specific refer to the same constructs that have previously been called Basic and Technical, or Common and Technical.

Figure 1.10. Hierarchical Relationships Among Functional Categories, Task Factors, and Task Constructs

Hands-On Performance Test

1. Safety/survival performance score
2. General (common) task performance score
3. Communication performance score
4. Vehicles performance score
5. MOS-specific task performance score

Job Knowledge Test

6. Safety/survival knowledge score
7. General (common) task knowledge score
8. Communication knowledge score
9. Identify targets knowledge score
10. Vehicles knowledge score
11. MOS-specific task knowledge score

Army-Wide Rating Scales

12. Overall effectiveness rating
13. Technical skill and effort composite
14. Personal discipline composite
15. Physical fitness/military bearing composite

MOS-Specific Rating Scales

16. Overall MOS composite

Combat Performance Prediction Scales

17. Overall Combat Prediction scale composite (available for males only)

Personnel File Form

18. Awards and Certificates
19. Disciplinary Actions (Articles 15 and Flag Actions)
20. Physical Readiness
21. M16 Qualification
22. Promotion Rate

Figure 1.11. Summary List of LVI Basic Criterion Scores

Table 1.8

LVI Sample Sizes for Performance Measures for Batch A MOS

MOS		N	Hands-On	Job Knowledge	Army-Wide Ratings	MOS Ratings	Combat Ratings	Personnel File	Combined Criteria[a]
11B	Infantryman	909	890	895	899	899	898	906	907
13B	Cannon Crewmember	916	773	810	897	897	897	916	916
19E	M60 Armor Crewman	249	243	248	241	241	241	249	249
19K	M1 Armor Crewman	824	749	812	782	778	782	819	825
31C	Single Channel Radio Operator	529	446	504	497	481	442	527	529
63B	Light-Wheel Vehicle Mechanic	752	624	723	728	719	666	750	752
71L	Administrative Specialist	678	641	664	634	626	199	675	678
88M	Motor Transport Operator	682	588	674	666	663	479	680	682
91A	Medical Specialist	824	794	798	807	797	670	818	824
95B	Military Police	452	444	446	451	450	366	452	452
Total		6,815	6,192	6,574	6,602	6,547	5,640	6,792	6,814

[a] Combined Criteria include Hands-On, Job Knowledge, Army-Wide Ratings, MOS Ratings, and Personnel File Form.

Table 1.9

LVI Sample Sizes for Performance Measures for Batch Z MOS

MOS		N	Job Knowledge	Army-Wide Ratings	Combat Ratings	Personnel File
12B	Combat Engineer	841	840	827	827	838
16S	MANPADS Crewman	472	471	468	468	472
27E	Tow/Dragon Repairer	90	90	89	84	90
29E	Comm.-Electronics Radio Repairer	112	111	106	101	111
51B	Carpentry/Masonry Specialist	213	212	193	190	212
54B	NBC Specialist	499	498	492	462	498
55B	Ammunition Specialist	279	279	269	243	279
67N	Utility Helicopter Repairer	197	194	193	192	197
76Y	Unit Supply Specialist	788	788	734	616	787
94B	Food Service Specialist	832	932	818	717	931
96B	Intelligence Analyst	128	128	122	103	128
Total		4,451	4,443	4,311	4,003	4,443

Table 1.10

LVI Combined Criteria Data: Percentage of Missing Data for Basic Scores by MOS

Criteria	11B	13B	19E	19K	31C	63B	71L	88M	91A	95B
Hands-On - Task Factors										
C - Communications	1.87	15.61	2.41	9.21	15.69	--	--	--	--	--
V - Vehicles	--	--	--	--	15.69	17.02	--	13.78	--	1.77
B - Basic Soldiering	1.87	15.61	2.41	9.21	15.69	17.02	5.46	13.78	3.64	1.77
I - Identify Targets	--	--	--	--	--	--	--	--	--	--
T - Technical	--	15.61	2.41	9.21	15.59	17.02	5.46	--	3.64	1.77
S - Safety/Survival	1.87	15.61	2.41	9.21	15.69	17.02	5.46	13.78	3.64	1.77
Job Knowledge - Task Factors										
C - Communications	2.65	12.01	.40	1.94	6.62	--	--	--	--	1.55
V - Vehicles	--	--	--	--	6.62	5.05	--	1.91	4.98	1.55
B - Basic Soldiering	2.65	12.01	.40	1.94	6.62	5.05	2.36	1.91	4.98	1.55
I - Identify Targets	2.65	12.01	.40	1.94	6.62	--	--	1.91	4.98	1.55
T - Technical	--	12.01	.40	1.94	6.62	5.05	2.36	--	4.98	1.55
S - Safety/Survival	2.65	12.01	.40	1.94	6.62	5.05	2.36	1.91	4.98	1.55
Army-Wide Ratings										
Overall Effectiveness	1.10	2.95	3.21	5.33	6.99	3.32	7.67	2.79	2.31	.22
Technical Skill and Effort	.88	2.07	3.21	5.21	6.05	3.19	6.93	2.35	2.06	.22
Personal Discipline	.88	2.07	3.21	5.21	6.05	3.19	6.93	2.35	2.06	.22
Physical Fitness/Bearing	.88	2.07	3.21	5.21	6.05	3.19	6.93	2.35	2.06	.22
MOS Ratings										
MOS Composite Rating	1.32	5.35	3.21	6.67	9.83	4.65	9.73	4.55	6.43	1.55
Personnel File Form										
Awards and Certificates	2.43	3.60	2.01	3.76	4.91	3.19	2.65	2.79	2.91	1.77
Articles 15 and Flag Actions	1.21	1.53	.00	1.82	1.51	1.06	1.18	.73	1.94	.44
Physical Readiness Score	4.63	5.46	3.21	5.21	9.45	11.44	9.00	9.09	6.55	5.31
M16 Qualification	2.65	4.04	29.32	18.30	2.65	3.19	1.77	2.93	3.88	3.98
Promotion Rate	1.76	1.86	.80	4.00	5.10	4.79	5.01	3.96	2.67	0.88

Note. -- indicates that the particular score was not calculated for that MOS.

Table 1.11

LVI Predictor Data: Amount of Missing Data for Paper-and-Pencil Scale Scores

Score	Not Missing	Missing
Assembling Objects - Number Correct	49,042	366
Map - Number Correct	49,047	361
Maze - Number Correct	49,052	356
Object Rotation - Number Correct	49,103	305
Orientation - Number Correct	49,072	336
Reasoning - Number Correct	49,103	305
JOB Scale 1 - Pride	46,525	2,883
JOB Scale 2 - Job Security/Comfort	46,634	2,774
JOB Scale 3 - Serving Others	46,295	3,113
JOB Scale 4 - Job Autonomy	46,037	3,371
JOB Scale 5 - Routine	45,975	3,433
JOB Scale 6 - Ambition	46,058	3,350
ABLE Scale 1 - Emotional Stability	44,264	5,144
ABLE Scale 2 - Self-Esteem	44,247	5,161
ABLE Scale 3 - Cooperativeness	44,258	5,150
ABLE Scale 4 - Conscientiousness	44,199	5,209
ABLE Scale 5 - Nondelinquency	44,228	5,180
ABLE Scale 6 - Traditional Values	44,190	5,218
ABLE Scale 7 - Work Orientation	44,260	5,148
ABLE Scale 8 - Internal Control	44,254	5,154
ABLE Scale 9 - Energy Level	44,217	5,191
ABLE Scale 10 - Dominance	44,246	5,162
ABLE Scale 11 - Physical Condition	44,264	5,144
AVOICE Scale 1 - Clerical/Administrative	45,477	3,931
AVOICE Scale 2 - Mechanics	45,941	3,467
AVOICE Scale 3 - Heavy Construction	45,851	3,557
AVOICE Scale 4 - Electronics	45,922	3,486
AVOICE Scale 5 - Combat	45,939	3,469
AVOICE Scale 6 - Medical Services	45,545	3,863
AVOICE Scale 7 - Rugged Individualism	45,944	3,464
AVOICE Scale 8 - Leadership/Guidance	45,508	3,900
AVOICE Scale 9 - Law Enforcement	45,958	3,450
AVOICE Scale 10 - Food Service Professional	45,916	3,492
AVOICE Scale 11 - Firearms Enthusiast	45,942	3,466
AVOICE Scale 12 - Science/Chemical	45,970	3,438
AVOICE Scale 13 - Drafting	45,976	3,432
AVOICE Scale 14 - Audiographics	45,452	3,956
AVOICE Scale 15 - Aesthetics	45,279	4,129
AVOICE Scale 16 - Computers	45,554	3,854
AVOICE Scale 17 - Food Service Employee	45,965	3,443
AVOICE Scale 18 - Mathematics	45,691	3,717
AVOICE Scale 19 - Electronic Communications	45,602	3,806
AVOICE Scale 20 - Warehousing/Shipping	45,963	3,445
AVOICE Scale 21 - Fire Protection	45,972	3,436
AVOICE Scale 22 - Vehicle Operator	45,971	3,437

Table 1.12

LVI Predictor Data: Amount of Missing Data for Computer-Administered Scale Scores

Score	Not Missing	Missing
Target Identification - Mean of Clipped Decision Time	38,401	513
Target Identification - Proportion Correct	38,404	510
Number Memory - Mean of Clipped Operation Means	38,324	590
Number Memory - Proportion Correct	38,353	561
Target Track 1 - Mean Log (Distance+1)	38,825	89
Target Track 2 - Mean Log (Distance+1)	38,793	121
Cannon Shoot - Mean Absolute Time Discrepancy	38,603	311
Target Shoot - Mean Log (Distance+1)	37,477	1,437
Mean of Median Movement Times across 5 tests	37,863	1,051
Simple Reaction Time - Median Decision Time	38,747	167
Simple Reaction Time - Proportion Correct	38,747	167
Choice Reaction Time - Median Decision Time	38,856	58
Choice Reaction Time - Proportion Correct	38,856	58
Perceptual Speed/Accuracy - Mean of Clipped Decision Time	38,703	211
Perceptual Speed/Accuracy - Proportion Correct	38,734	180
Short-Term Memory - Mean of Clipped Decision Time	38,483	431
Short-Term Memory - Proportion Correct	38,490	424

An imputation procedure known as PROC IMPUTE (Wise & McLaughlin, 1980) was developed that used existing data to estimate values for missing data. This procedure was also used in the CVI analyses (Wise, McHenry, & Young, 1986). The decision rules used in the CVI analyses were replicated in the LVI analyses as closely as possible.

PROC IMPUTE uses regression estimates to predict missing values. Each missing value is predicted from other values for the subject in question so that individual differences are retained. The regression coefficient and intercept vary from item to item so that differences in item difficulty are also reflected in the predicted values. PROC IMPUTE also adds a random variable with variance equal to the error of estimate for predicting the missing value.

The results of the imputation were examined at two levels. First, after each PROC IMPUTE run, the program output was inspected. Second, the pre-imputed and the post-imputed data sets were compared for each MOS (a) after

the hands-on score level imputation, and (b) after the criterion construct level imputation.

The means and variances of the pre- and post-imputation results for the hands-on data for each MOS were found to be virtually identical. Imputation also made virtually no difference in the magnitude of the intercorrelations among the criterion scores that were used to create the performance factor scores in the validation analyses. These results are similar to those obtained earlier from the CVI imputation (Wise et al., 1986).

Development of the LVI First-Tour Performance Model

A latent factor model of first-tour performance, developed using data from the Project A Concurrent Validation (CVI) sample, has been described by Campbell, McHenry, and Wise (1990). This model included the now familiar five performance factors--Core Technical Proficiency (CTP), General Soldiering Proficiency (GSP), Effort and Leadership (ELS), Maintaining Personal Discipline (MPD), and Physical Fitness and Military Bearing (PFB)-- and two measurement method factors, a Ratings method factor and a Paper-and-Pencil Test method factor. During year two, the CVI model was subjected to a confirmatory analysis, using first-tour performance data collected from the Longitudinal Validation (LVI) sample. Additionally, comparative analyses aimed at evaluating more parsimonious models of first-tour performance were carried out.

An earlier section summarized how each of the major sets of performance measures was reduced from a large number of item, task, or individual scale scores to a smaller set of factor or category scores. The results of this first level of aggregation have been referred to as the "basic" array of criterion scores, summarized in Figure 1.11. These included the scores that were used in the modeling analyses described below.

Altogether, the LVI first-tour performance measures were reduced to 20 basic scores. However, because MOS differ in their task content, not all 20 variables were scored in each MOS, and there was some slight variation in the number of variables used in the subsequent analyses.

To test the fit of the different models to the LVI data, confirmatory factor-analytic techniques were applied to each MOS individually, using LISREL 7 (Jöreskog & Sörbom, 1989). The first alternative five-factor model was developed using CVI data. After the fit of the five-factor model was assessed in each MOS, four reduced models (all nested within the five-factor model) were examined. Finally, as had been done in the original CVI analyses, the five-factor model was applied to the Batch A MOS simultaneously (using LISREL's multigroups option). The fit statistics (e.g., root mean-square residuals [RMSRs]) of the five-factor model for each MOS in the LVI and CVI samples were very similar. In fact, for three of the MOS (11B, 13B, and 71L), the RMSRs for the LVI data were smaller than those for the CVI data. These results indicate that the model developed using the CVI data does fit the LVI data quite well.

Four reduced models were also examined using the LVI data. For the <u>four-factor model</u>, the Core Technical Proficiency and General Soldiering Proficiency performance factors were collapsed into a single "can do" performance factor. The <u>three-factor model</u> retained the "can do" performance

factor of the four-factor model, but also collapsed the Effort and Leadership and Maintaining Personal Discipline performance factors into a "will do" performance factor. For the two-factor model, the "can do" performance factor was retained; however, the Physical Fitness and Military Bearing performance factor became part of the "will do" performance factor. Finally, for the one-factor model, the "can do" and "will do" performance factors, or equivalently, the five original performance factors, were collapsed into a single performance factor.

The chi-square statistics and RMSRs, respectively, for the four reduced models, as well as for the five-factor model, indicate that the four- and five-factor models fit the LVI data well, while the one-, two-, and three-factor models fit less well. The results also indicated that the parameter estimates for the five-factor model were generally similar across the 10 MOS. The final step was to determine whether the variation in some of these parameters could be attributed to sampling variation. To do this (as described earlier), the following were specified to be invariant across jobs: (a) the correlations among performance factors, (b) the loadings of all the Army-wide measures on the performance factors and on the rating method factor, (c) the loadings of the MOS-specific score on the rating method factor, and (d) the uniqueness coefficients for the Army-wide measures.

The results indicated that the fit of the five-factor model is not as good when the parameters listed above are constrained to be equal across the 10 jobs. Still, the root mean-square residuals associated with the across-MOS model are not substantially greater than those for the within-job analyses. (The average RMSR for the across-MOS model is .0676; the average for the within-MOS models is .0585.)

To create criterion construct scores for use in validation analyses, the scoring procedures were based on the five-factor model. Although the four-factor model has the advantage of greater parsimony, the five-factor model offered the advantage of corresponding to the criterion constructs generated in the CVI validation analyses. Table 1.13 shows the mapping of the basic scores on the five performance factors. As with the CVI data, five residual scores, corresponding to the five criterion constructs, were also created.

The five "raw" criterion construct scores, the five residual criterion construct scores, the total rating and job knowledge scores, and the total score derived from the hands-on test were used to generate a 13 x 13 matrix of criterion intercorrelations for each MOS in Batch A. The averages of these correlations are reported in Table 1.14. These results are very similar to the correlations that were reported by Campbell et al. (1990) for the CVI sample.

Basic Validation Results for the LVI Sample

The LVI validation results were based on two different sample editing strategies. The first required complete data for all predictor composites, as well as for the ASVAB, and for each performance factor; this sample is referred to as the "listwise deletion" sample. In the alternative strategy, called setwise deletion, a separate validation sample was identified for each set of predictors in the Experimental Battery.

Table 1.13

Mapping of LVII Performance Measures Onto Latent Performance Factors

	Performance Factors					Method Factors	
Criterion Score[a]	Core Technical Proficiency	General Soldiering Proficiency	Effort and Leadership	Maintaining Personal Discipline	Physical Fitness/ Military Bearing	Written Know-ledge Tests	Rating Scales
HO Technical	X						
HO Communication		X					
HO Vehicles		X					
HO General Soldier		X					
HO Safety/Survival		X					
JK Technical	X					X	
JK Communication		X				X	
JK Vehicles		X				X	
JK General Soldier		X				X	
JK ID Threat/Target		X				X	
JK Safety/Survival		X				X	
AWB Skill/Effort Composite			X				X
AWB Discipline Composite				X			X
AWB Fitness Composite					X		X
AWB Overall Composite			X	X			X
MOS Rating Composite			X			X	
PFF Awards/Certificates			X				
PFF Physical Readiness					X		
PFF Articles 15/Flags				X			
PFF Promotion Rate				X			

[a] AWB = Army-Wide Rating Scales; HO = Hands-On; JK = Job Knowledge; PFF = Personnel File Form.

Table 1.14

Mean Intercorrelations Among 13 Summary Criterion Scores for the Batch A MOS in the LVI Sample

Summary Criterion Score[a]	CTP Raw	GSP Raw	ELS Raw	MPD Raw	PFB Raw	CTP Res	GSP Res	ELS Res	MPD Res	PFB Res	PRT	HOT	JKT
CTP (raw)	1.00												
GSP (raw)	.57	1.00											
ELS (raw)	.25	.26	1.00										
MPD (raw)	.16	.18	.58	1.00									
PFB (raw)	.06	.06	.48	.36	1.00								
CTP (residual)	.88	.41	.30	.20	.07	1.00							
GSP (residual)	.40	.88	.32	.23	.06	.45	1.00						
ELS (residual)	.41	.42	.70	.43	.26	.40	.42	1.00					
MPD (residual)	.20	.22	.28	.88	.17	.20	.23	.46	1.00				
PFB (residual)	.07	.07	.20	.21	.90	.04	.03	.29	.21	1.00			
Perf. Rating Total	.22	.24	.88	.72	.58	.27	.28	.40	.35	.24	1.00		
Hands-On Total	.72	.76	.26	.15	.08	.81	.85	.41	.18	.09	.23	1.00	
Job Knowledge Total	.74	.80	.25	.19	.04	.40	.46	.40	.23	.04	.22	.47	1.00

[a] CTP = Core Technical Proficiency; GSP = General Soldiering Proficiency; ELS = Effort and Leadership; MPD = Maintaining Personal Discipline; PFB = Physical Fitness and Military Bearing.

The number of soldiers with complete predictor and criterion data in each MOS is reported in Table 1.15 for both the CVI and LVI data sets.

Table 1.15

Soldiers in CVI and LVI Data Sets With Complete Predictor and First-Tour Criterion Data by MOS

MOS		CVI	LVI (Listwise Deletion Sample)
11B	Infantryman	491	235
13B	Cannon Crewmember	464	553
19E[a]	M60 Armor Crewman	394	73
19K	M1 Armor Crewman	---	446
31C	Single Channel Radio Operator	289	172
63B	Light-Wheel Vehicle Mechanic	478	406
71L	Administrative Specialist	427	252
88M	Motor Transport Operator	507	221
91A	Medical Specialist	392	535
95B	Military Police	597	270
Total		4,039	3,163

[a] MOS 19E not included in LVI validity analyses.

The analysis procedure consisted of the following major steps:

A) Using the listwise deletion sample, multiple correlations between each set of predictor scores and the five substantive factor scores, their five residual factor scores, the two method factor scores, and the total scores from the hands-on and job knowledge tests were computed separately by MOS and then averaged.

B) Using the listwise deletion sample, incremental validities for each set of Experimental Battery predictors (e.g., AVOICE composites or computer composites) over the four ASVAB factor composites were computed against the same criteria used to compute the validities in Step A. Once again, the results were computed separately by MOS and then averaged.

C) Using the setwise deletion samples, multiple correlations and incremental validities (over the four ASVAB factor composites) between each set of Experimental Battery predictors and the criteria used in the first two steps were computed separately by MOS and then averaged. All results to this point were corrected for range restriction and adjusted for shrinkage using the Rozeboom formula.

D) Finally, once again using the listwise deletion sample, multiple correlations and incremental validities (over the four ASVAB factors) were computed for each set of predictors in the Experimental Battery, this time adjusting the results for shrinkage with the Claudy (1978) instead of the Rozeboom formula. This step was conducted to allow comparisons between the first-tour validity results associated with the longitudinal sample and those that had been reported for the concurrent sample (for which only the Claudy formula was used, e.g., McHenry, Hough, Toquam, Hanson, & Ashworth, 1990).

Multiple Correlations and Incremental Validities Based on Listwise Deletion Samples

Multiple correlations for the four ASVAB factor composites, the single spatial composite, the eight computer composites, the three JOB composites, the seven ABLE composites, and the eight AVOICE composites are reported in Table 1.16.

Incremental validity results for the Experimental Battery predictors over the ASVAB factors are reported in Table 1.17. The results indicate that the spatial composite added slightly to the prediction of the raw and residual Core Technical and General Soldiering performance factors, as well as to the written method factor and the hands-on and job knowledge total scores. They also show that the seven ABLE composites contributed substantially to the prediction of the raw and residual Personal Discipline and Physical Fitness performance factors.

Multiple Correlations and Incremental Validities Based on the Setwise Deletion Samples

Multiple correlations for the spatial composite, the eight computer composites, the three JOB composites, the seven ABLE composites, and the eight AVOICE composites based on the setwise deletion samples described above are reported in Table 1.18. These multiple correlations were very similar to those computed with the listwise sample. However, there was a consistent difference between the two sets of results; specifically, the multiple correlations based on the setwise samples were generally one to three validity points higher.

Incremental validity results associated with the setwise deletion samples can be found in Table 1.19. The incremental validity results based on the setwise samples were practically identical to those based on the listwise sample. Again, the primary difference between the two sets of results was that the level of validities was sometimes one or two points lower for the listwise sample than for the setwise samples.

Table 1.16

Mean of Multiple Correlations Computed Within Job for LVI Listwise Deletion Samples for ASVAB Factors, Spatial, Computer, JOB, ABLE Composites, and AVOICE

Criterion[a]	No. of MOS[b]	ASVAB Factors [4]	Spatial [1]	Computer [8]	JOB [3]	ABLE Comp. [7]	AVOICE [8]
CTP (Raw)	9	62 (13)	57 (11)	47 (16)	29 (13)	21 (09)	38 (08)
GSP (Raw)	8	66 (07)	64 (06)	55 (08)	29 (13)	23 (14)	37 (07)
ELS (Raw)	9	37 (12)	32 (08)	29 (15)	18 (14)	13 (11)	17 (15)
MPD (Raw)	9	17 (13)	14 (11)	10 (16)	06 (13)	14 (11)	05 (10)
PFB (Raw)	9	16 (06)	10 (04)	07 (07)	06 (06)	27 (07)	05 (09)
CTP (Res)	9	46 (17)	42 (15)	29 (22)	17 (12)	08 (11)	28 (12)
GSP (Res)	8	51 (10)	51 (08)	41 (10)	18 (11)	12 (12)	26 (09)
ELS (Res)	9	46 (18)	41 (13)	37 (20)	23 (15)	21 (15)	24 (16)
MPD (Res)	9	18 (13)	14 (12)	08 (16)	07 (11)	13 (11)	06 (10)
PFB (Res)	9	20 (10)	12 (08)	09 (11)	07 (06)	28 (10)	09 (11)
Written	9	54 (13)	49 (12)	43 (18)	29 (16)	23 (12)	29 (14)
Ratings	9	12 (09)	09 (07)	07 (09)	06 (09)	03 (05)	02 (07)
HO-Total	9	50 (14)	48 (11)	38 (15)	18 (13)	11 (11)	28 (09)
JK-Total	9	71 (08)	65 (07)	58 (10)	36 (14)	31 (08)	41 (08)

Note. Corrected for range restriction, and adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] CTP = Core Technical Proficiency; GSP = General Soldiering Proficiency; ELS = Effort and Leadership; MPD = Maintaining Personal Discipline; PFB = Physical Fitness and Military Bearing; HO = Hands-On; JK = Job Knowledge.

[b] Number of MOS for which validities were computed.

Table 1.17

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVI Listwise Deletion Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE

Criterion	No. of MOS[a]	ASVAB Factors (A4) [4]	A4+ Spatial [5]	A4+ Computer [12]	A4+ JOB [7]	A4+ ABLE Comp. [11]	A4+ AVOICE [12]
CTP (Raw)	9	*62* (13)	**63** (13)	61 (14)	61 (13)	61 (13)	62 (13)
GSP (Raw)	8	*66* (07)	**68** (07)	66 (07)	66 (07)	66 (07)	66 (07)
ELS (Raw)	9	*37* (12)	36 (13)	35 (13)	36 (13)	34 (17)	33 (16)
MPD (Raw)	9	*17* (13)	16 (14)	16 (15)	14 (15)	**23** (14)	10 (15)
PFB (Raw)	9	*16* (06)	13 (08)	09 (08)	**17** (08)	**30** (06)	12 (10)
CTP (Res)	9	*46* (17)	**47** (17)	44 (18)	45 (18)	43 (19)	46 (19)
GSP (Res)	8	*51* (10)	**53** (09)	51 (10)	50 (10)	50 (10)	50 (10)
ELS (Res)	9	*46* (18)	**47** (18)	44 (21)	45 (21)	45 (22)	44 (21)
MPD (Res)	9	*18* (13)	15 (14)	15 (14)	14 (14)	**22** (14)	12 (13)
PFB (Res)	9	*20* (10)	18 (12)	13 (11)	20 (11)	**34** (10)	18 (13)
Written	9	*54* (13)	**55** (13)	51 (18)	54 (13)	54 (12)	52 (17)
Ratings	9	*12* (09)	11 (08)	09 (10)	09 (10)	09 (08)	05 (08)
HO-Total	9	*50* (14)	**52** (13)	49 (14)	49 (15)	48 (14)	49 (15)
JK-Total	9	*71* (08)	**72** (08)	71 (09)	71 (08)	71 (08)	71 (08)

Note. Corrected for range restriction, and adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Multiple Rs for ASVAB Factors alone are in italics. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 1.18

Mean of Multiple Correlations Computed Within Job for LVI Setwise Deletion Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE

Criterion	No. of MOS[a]	Spatial [1]	Computer [8]	JOB [3]	ABLE Composites [7]	AVOICE [8]
CTP (Raw)	9	58 (11)	49 (16)	31 (13)	21 (09)	39 (07)
GSP (Raw)	8	65 (06)	55 (08)	32 (13)	24 (14)	38 (07)
ELS (Raw)	9	33 (08)	30 (15)	19 (14)	12 (11)	20 (12)
MPD (Raw)	9	14 (11)	10 (16)	06 (13)	15 (11)	05 (11)
PFB (Raw)	9	08 (04)	13 (07)	07 (06)	28 (07)	09 (09)
CTP (Res)	9	43 (15)	31 (22)	17 (12)	10 (11)	29 (09)
GSP (Res)	8	51 (08)	40 (10)	21 (11)	14 (12)	28 (09)
ELS (Res)	9	41 (13)	36 (20)	24 (15)	21 (15)	26 (06)
MPD (Res)	9	13 (12)	10 (16)	06 (11)	15 (11)	07 (13)
PFB (Res)	9	11 (08)	10 (11)	09 (06)	30 (10)	12 (10)
Written	9	51 (11)	46 (16)	31 (17)	25 (11)	32 (15)
Ratings	9	09 (08)	09 (09)	07 (08)	04 (06)	03 (07)
HO-Total	9	50 (11)	38 (15)	20 (13)	13 (11)	30 (07)
JK-Total	9	66 (07)	60 (10)	38 (14)	30 (08)	43 (08)

Note. Corrected for range restriction and adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 1.19

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVI Setwise Deletion Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE

Criterion	No. of MOS[a]	ASVAB Factors (A4) + Spatial [5]	A4+ Computer [12]	A4+ JOB [7]	A4+ ABLE Composites [11]	A4+ AVOICE [12]
CTP (Raw)	9	**64** (10)	61 (11)	63 (11)	61 (12)	64 (11)
GSP (Raw)	8	**69** (06)	**66** (07)	67 (07)	66 (08)	66 (07)
ELS (Raw)	9	37 (10)	36 (14)	37 (11)	36 (13)	36 (11)
MPD (Raw)	9	15 (13)	15 (15)	12 (13)	**24** (13)	11 (14)
PFB (Raw)	9	15 (08)	17 (05)	**17** (07)	**32** (04)	15 (10)
CTP (Res)	9	**48** (12)	45 (14)	46 (14)	45 (14)	47 (14)
GSP (Res)	8	**54** (06)	50 (08)	51 (08)	50 (07)	50 (07)
ELS (Res)	9	47 (12)	43 (20)	46 (15)	46 (15)	46 (14)
MPD (Res)	9	14 (13)	13 (15)	13 (13)	**22** (12)	11 (14)
PFB (Res)	9	20 (11)	18 (11)	20 (10)	**36** (08)	21 (11)
Written	9	**57** (13)	53 (17)	58 (12)	55 (13)	54 (18)
Ratings	9	10 (09)	**11** (11)	11 (09)	**11** (07)	06 (09)
HO-Total	9	**53** (09)	49 (11)	50 (12)	49 (11)	50 (11)
JK-Total	9	**73** (08)	71 (09)	72 (08)	71 (09)	71 (09)

Note. Corrected for range restriction and adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone. Decimals omitted.

[a] Number of MOS for which validities were computed.

Comparison of Validity Research in LVI and CVI Samples

The final set of results concern the comparison between the validity estimates associated with the longitudinal data (i.e., LVI) and those reported for the concurrent validation data (CVI). Table 1.20 reports the multiple correlations for the ASVAB factors and each set of experimental predictors as computed for the listwise sample in both data sets.

The results in Table 1.20 demonstrate that the patterns and levels of validities are very similar across the two sets of analyses. Still, there are some differences worth pointing out. Specifically, in comparison to the results of the CVI analyses: (a) The LVI validities of the "cognitive" predictors (i.e., ASVAB, spatial, computer) for predicting the "will do" performance factors (ELS, MPD, and PFB) are higher; (b) the LVI validities of the ABLE composites for predicting the "will do" performance factors are somewhat lower; and (c) the LVI validities of the AVOICE composites for predicting the "can do" performance factors (CTP and GSP) are higher.

Table 1.20

Comparison of Mean Multiple Correlations Computed Within Job for LVI and CVI Listwise Deletion Samples for ASVAB Factors, Spatial, Computer, JOB, ABLE Composites, and AVOICE

Criterion	No. of MOS[a]	ASVAB Factors		Spatial		Computer		JOB		ABLE Comp.		AVOICE	
		LV [4]	CV [4]	LV [1]	CV [1]	LV [8]	CV [6]	LV [3]	CV [3]	LV [7]	CV [4]	LV [8]	CV [6]
CTP (Raw)	9	63	63	57	56	50	53	31	29	27	26	41	35
GSP (Raw)	8	67	65	64	63	57	57	32	30	29	25	40	34
ELS (Raw)	9	39	31	32	25	34	26	22	19	20	33	25	24
MPD (Raw)	9	22	16	14	12	15	12	11	11	22	32	11	13
PFB (Raw)	9	21	20	10	10	17	11	12	11	31	37	15	12
CTP (Res)	9	48	47	42	37	35	37	20	21	18	22	33	28
GSP (Res)	8	53	49	51	48	44	41	22	22	19	21	31	26
ELS (Res)	9	48	46	41	41	40	38	25	27	26	31	29	32
MPD (Res)	9	23	19	14	15	14	13	12	10	21	28	13	15
PFB (Res)	9	24	21	12	11	17	14	11	10	32	35	16	14
Written	9	56	62	49	55	47	54	31	28	29	21	33	32
Ratings	9	16	15	09	07	17	08	10	08	09	18	09	09

Note. Corrected for range restriction and adjusted for shrinkage (Claudy formula). Numbers in brackets are the numbers of predictor scores entering prediction equations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Further Exploration of ELS and ABLE

As shown in the date reported above, the largest difference between the CVI and LVI validation results was in the prediction of the Effort and Leadership (ELS) performance factors with the ABLE basic scores. Corrected for restriction of range and for shrinkage, the validity of the four ABLE composite scores in CVI was .33 for ELS and the validity of the seven ABLE factor scores in LVI was .20. When cast against the variability in results across studies in the extant literature, such a difference may not seem all that large or very unusual. However, since the obtained results from CVI, CVII, and LVI have been so consistent, in terms of the expected convergent and divergent results, we subjected this particular difference to a series of additional analyses in an attempt to determine the reason for the discrepancy.

First, the discrepancy does not seem to arise from any general deterioration in the measurement properties of either the ABLE or the ELS composite in the LVI sample. For example, while the correlation of the ABLE with ELS and MPD went down, the ABLE's correlations with CTP and GSP went up slightly. Similarly, a decrease in the validity with which ELS was predicted was characteristic only of the ABLE. The validities of the cognitive measures, the JOB, and AVOICE for predicting ELS actually increased by varying amounts. Consequently, the decrease in validity seems to be specific to the ABLE/ELS correlation and, to a lesser extent, the ABLE/MPD correlation.

The followup analyses were also able to rule out two possible additional sources of the CVI/LVI validity differences. First, differences in the composition and number of ABLE basic scores from CVI to LVI did not account for the differences in patterns of validity. Second, differences in the composition of the Effort/Leadership factor score from CVI to LVI did not account for differences in validity.

Rather, the somewhat lower correlation of ABLE with Effort/Leadership in LVI seems due to the joint effects of two influences. First, the determinants of ELS scores seem to favor ability slightly more and motivation slightly less in LVI versus CVI, perhaps because their true score variances were different across the two cohorts. Second, the greater influence of the social desirability response tendency in LVI seems to produce more positive manifold (i.e., higher intercorrelations for the LVI ABLE basic scores), as contrasted with CVI. This could also lower the correlation of the regression-weighted ABLE composite with ELS, whereas it might not have the same effect with the Core Technical and General Soldiering factors.

Yet another component of the explanation is the negative correlation between the Social Desirability scale and AFQT. AFQT and Social Desirability correlated -.22 in the CV sample and -.20 in the LV sample. This would tend to lower the correlation between ABLE and ELS if the correlations between ABLE and ASVAB and between ASVAB and ELS were positive, which they were.

Summary of LVI Validation

Generally speaking, the ASVAB was the best predictor of performance. However, the composite of spatial tests provided a small amount of incremental validity for the "can do" criteria (1-3 points), and the ABLE provided larger increments (7-20 points) for two of the three "will do" criteria (Maintaining Personal Discipline, and Physical Fitness and Bearing). Estimates of incremental

validity were somewhat higher when the results were not corrected for range restriction.

With regard to ASVAB scoring options, results indicate a very slight edge for using multiple regression equations based on the four ASVAB unit- weighted factor scores. In the test of ABLE scoring options, the method using factor scores computed from a subset of all the ABLE items (ABLE-114) proved to have consistently slightly higher validities.

Perhaps the most interesting finding is derived from the comparisons between the Longitudinal Validation results and those from the Concurrent Validation. Generally speaking, the pattern and level of the validity coefficients were highly similar across the two samples. The correlation between the CV and LV coefficients in Table 1.20 was .962 and the root mean-squared difference between the two sets of coefficients was .046. However, the correlation is not 1.00. As noted above, the longitudinal validities were higher for cognitive predictors against "will do" criteria and lower for ABLE composites against "will do" criteria. Some of the possible explanations for those differences include changes in the nature of predictor scores when administered in a longitudinal versus concurrent design, changes in criterion or predictor scores due to cohort differences, and changes in the true relationship between abilities and performance as persons gain more experience and training in an organization and job. These and other possible explanations will be explored in future analyses.

Results of the Concurrent Sample Second-Tour Validation (CVII)

The CVII validation results are based on the CVII sample, which was assessed on the criterion measures of second-tour performance at the same time that the LVI performance data were collected from the first-tour longitudinal sample. The predictor set is limited to ASVAB and ABLE because only a small proportion (approximately 12%) of the CVII sample had been assessed with the Experimental Predictor Battery. ASVAB scores, taken 5-6 years earlier, were available from the Enlisted Master File. The ABLE was administered concurrently during the CVII data collection to approximately 45 percent of the total sample (i.e., those individuals who had no peers in the sample to rate and thus had time to take the ABLE). Everyone in the sample was assessed on the full set of second-tour performance measures. By design, the MOS in the CVII sample were limited to the MOS in Batch A. Because of the generally small samples for individual MOS, results for most analyses are reported for the combined sample.

The CVII data collection and data presentation are described in the first annual report for Building the Career Force (Campbell & Zook, 1990; see Chapters 5 and 6). After final editing, the total N for CVII was 1,053. The total sample was distributed across the Batch A MOS as shown in Table 1.21.

Because of some missing data, the sample sizes varied depending on the specific analysis being reported. For example, for the reasons cited above, ABLE scores were available for only 477 individuals. All the analyses that require a common covariance matrix for ABLE and ASVAB were based on this reduced sample.

The development of the CVII performance measures, and the analysis and modeling of CVII performance, all have been described previously (Campbell & Zook, 1990) and are summarized in a previous section of the present chapter. The solution that yielded the best fit consisted of six substantive factors and two

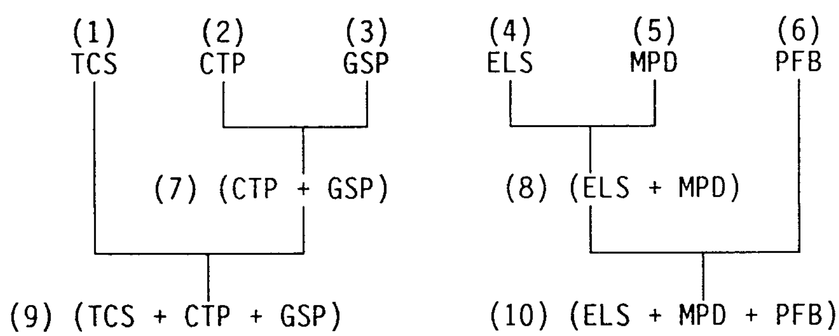
Table 1.21

CVII Sample Sizes by MOS

MOS	N
11B Infantryman	127
13B Cannon Crewmember	162
19E M60 Armor Crewman	33
19K M1 Armor Crewman	10
31C Single Channel Radio Operator	103
63B Light-Wheel Vehicle Mechanic	116
71L Administrative Specialist	112
88M Motor Transport Operator	144
91A Medical Specialist	146
95B Military Police	141
Total	1,053

methods factors. The two methods factors were defined to be orthogonal to the substantive factors, but the correlations among the substantive factors were not so constrained. The six substantive factors and two methods factors, and the variables that are scored on each, were shown in Figure 1.9.

The complete basic validation analyses utilized a total of 10 scores for the performance factors, as shown below.



TCS = Training/Counseling Subordinates; CTP = Core Technical Proficiency; GSP = General Soldiering Proficiency; ELS = Effort/Leadership; MPD = Maintaining Personal Discipline; PFB = Physical Fitness and Bearing

That is, all 10 scores were used as criterion measures. All higher order composite scores were obtained by standardizing the component scores and then taking the simple sum.

Procedure

The CVII validation analysis procedure consisted of the following steps.

(1) The ASVAB and ABLE were correlated with the six performance factor scores, their five residual scores (there was no residual for TCS), the higher order factor composites, the two methods factor scores, and the total score from the hands-on tests, the job knowledge tests, and the Situational Judgment Test. ASVAB was represented by the AFQT, a regression-weighted composite of the four factors, and a regression-weighted composite of the nine subtests. ABLE was represented by the three alternative sets of scores described previously. Both corrected (for multivariate restriction of range) and uncorrected estimates were computed, and both regression weights and unit weights (applied to standardized scores) were used. When multiple regression weights were used, the Rozeboom correction (Rozeboom, 1978) was used to account for the fitting of error.

(2) As in CVI, incremental validities for the ABLE composites over the ASVAB composites were also computed against each criterion score.

(3) A hierarchical regression analysis, stopping at six predictors, was run against each performance factor, factor composite, and individual criterion score (i.e., hands-on, job knowledge, and Situational Judgment Test).

(4) A hierarchical regression analysis was also carried out on selected criterion variables for the combined samples from three MOS clusters. The clusters were based on the results of an MOS clustering within the Synthetic Validation Project (Wise, Peterson, Hoffman, Campbell, & Arabian, 1991) and on the results of the validity generalization analysis for the Batch A MOS in the CVI sample (Wise, McHenry, & Campbell, 1990).

(5) The final step consisted of using the optimal six variable equations from the hierarchical regression analyses described above to develop a picture of the degree of differential prediction across performance factors and across the three MOS clusters.

Results

The basic multiple correlations for ASVAB (four factors vs. nine subtests) and ABLE (seven theoretically based composites vs. seven "purified" empirical factors) are given in Table 1.22. Several things are worth noting. ASVAB, taken at time of entry, is still a highly valid predictor of Core Technical and General Soldiering Proficiency and has respectable validity for Effort/Leadership. For ASVAB, the four factors and the nine subtests provide virtually the same level of predictive accuracy. However, for ABLE the reduced factor scores (114 items) are consistently the best predictor set. ABLE predicts Effort/Leadership and Physical Fitness very well and has reasonable correlations with General Soldiering and Training/Counseling.

Table 1.22

Multiple Correlations for ASVAB Factors, ASVAB Subtests, ABLE Composites, and ABLE-114 Scores Against 19 CVII Criterion Variables (All MOS), With Unit Weights

Variable	ASVAB Factors [4]	ASVAB Subtests [9]	ABLE Composites [7]	ABLE-114 [7]
Core Technical (Raw)	43 (42)	43 (43)	15 (14)	20 (15)
General Soldiering (Raw)	56 (54)	57 (55)	22 (16)	26 (18)
Effort/Leadership (Raw)	38 (38)	39 (38)	37 (32)	41 (32)
Personal Discipline (Raw)	00 (11)	00 (11)	20 (21)	18 (22)
Physical Fitness (Raw)	13 (16)	06 (16)	32 (23)	34 (21)
Training/Counseling (Raw)	06 (13)	00 (12)	27 (19)	23 (18)
Core Technical (Res)	29 (29)	28 (30)	00 (12)	07 (13)
General Soldiering (Res)	42 (42)	43 (42)	14 (15)	18 (16)
Effort/Leadership (Res)	25 (26)	27 (25)	38 (31)	41 (30)
Personal Discipline (Res)	00 (09)	00 (09)	16 (20)	15 (19)
Physical Fitness (Res)	16 (20)	09 (20)	34 (21)	35 (18)
ELS - No Situational Judgment	24 (22)	23 (22)	34 (31)	38 (30)
Criterion Composite CTP/GSP	57 (55)	58 (56)	22 (17)	27 (19)
Criterion Composite ELS/MPD	29 (30)	29 (29)	34 (32)	37 (32)
Criterion Factor 1 CTP+GSP+TCS	50 (50)	50 (50)	29 (22)	32 (23)
Criterion Factor 2 ELS+MPD+PFB	14 (16)	12 (15)	34 (35)	35 (34)
Hands-On Average	39 (40)	38 (40)	12 (12)	18 (13)
Job Knowledge Total	59 (56)	59 (57)	25 (14)	28 (16)
Situational Judgment	42 (43)	42 (43)	27 (20)	31 (21)

Note. N = 412. Adjusted (Rozeboom formula). Numbers in brackets are the number of predictor scores entering prediction equations. Validities of unit-weighted composites are in parentheses. Decimals omitted.

In general, after adjustments, regression weights and unit weights for ASVAB yield about the same level of validity. However, regression weights are somewhat better than unit weights for the seven empirical ABLE factors. There is not as much positive manifold among the ABLE factors as there is among the ASVAB subtests.

Table 1.23 contains the same type of incremental analyses that were done in CVI (Campbell & Zook, 1991). ABLE does not add to the prediction of Core Technical and General Soldiering Proficiency, but it adds about the same amount to the prediction of Effort/Leadership as it did in CVI. However, the overall level of prediction for ELS is <u>higher</u> in CVII than it was in CVI ($\underline{R}$ = .50 vs. .43).

Table 1.23

Multiple Correlations for ASVAB Factors Plus ABLE Composites and Plus ABLE-114 Scores, and for ASVAB Subtests Plus ABLE Composites and Plus ABLE-114 Scores Against 19 CVII Criterion Variables, All MOS

Variable	4 ASVAB Factors + 7 ABLE Comp (K=11)	4 ASVAB Factors + 7 ABLE-114 (K=11)	9 ASVAB Subtests + 7 ABLE Comp (K=16)	9 ASVAB Subtests + 7 ABLE-114 (K=16)
Core Technical (Raw)	.42	.43	.42	.43
General Soldiering (Raw)	.56	.57	.58	.58
Effort/Leadership (Raw)	.49	.49	.49	.50
Personal Discipline (Raw)	.16	.13	.09	.03
Physical Fitness (Raw)	.34	.35	.32	.33
Training/Counseling (Raw)	.26	.20	.24	.17
Core Technical (Res)	.24	.26	.24	.25
General Soldiering (Res)	.42	.42	.44	.44
Effort/Leadership (Res)	.43	.43	.43	.43
Personal Discipline (Res)	.09	.07	.00	.00
Physical Fitness (Res)	.36	.37	.34	.34
ELS - No Situational Judgment	.39	.41	.38	.41
Criterion Composite CTP/GSP	.57	.57	.58	.58
Criterion Composite ELS/MPD	.40	.40	.40	.40
Criterion Factor 1: CTP+GSP+TCS	.54	.54	.54	.54
Criterion Factor 2: ELS+MPD+PFB	.35	.35	.34	.34
Hands-On Average	.37	.37	.37	.37
Job Knowledge Total	.60	.60	.60	.60
Situational Judgment	.45	.44	.45	.44

Note. N = 412. Corrected for range restriction and adjusted (Rozeboom formula).

The hierarchical procedure asked for the optimal six-variable equation. For any specific criterion measure the first four variables were never all from ASVAB or all from ABLE. It appears that ABLE, most frequently the Dependability scale, does play a role in predicting CTP and GSP. This contribution is masked when the non-hierarchical procedure is used.

Generalizability

A descriptive picture of the generalizability of prediction equations across performance factors (for the combined sample) is shown in Table 1.24. All entries are multiple correlations and the diagonals represent estimates based on optimal weights. Estimates of what happens when less than optimal weights are used to predict the same criterion are obtained by looking across the rows. Estimates of what happens when a particular set of weights is applied to other criterion measures or other MOS are obtained by looking down the columns. All estimates are based on the corrected covariance matrix. The diagonals are adjusted for shrinkage using the Rozeboom formula with k = 6. The off-diagonals are not adjusted because the weights were not computed against that particular dependent variable.

Table 1.24

Multiple Correlations for 10 Sets of Criterion Composite Weights, All MOS

	Raw CTP Weights	Raw GSP Weights	Raw ELS Weights	Raw MPD Weights	Raw PFB Weights	Raw TCS Weights	CTP+GSP Weights	ELS+MPD Weights	Criterion 1 Weights	Criterion 2 Weights
Core Technical (Raw CTP)	**.451** (.429)	.436	.331	.165	.193	.195	.446	.298	.434	.154
General Soldiering (Raw GSP)	.553	**.571** (.557)	.422	.173	.232	.288	.568	.367	.561	.192
Effort/ Leadership (Raw ELS)	.368	.370	**.500** (.482)	.375	.046	.358	.372	.489	.404	.422
Personal Discipline (Raw MPD)	.083	.069	.169	**.226** (.169)	.057	.130	.075	.197	.094	.188
Physical Fitness (Raw PFB)	.171	.163	.037	.100	**.401** (.375)	.055	.168	.059	.135	.235
Training/ Counseling (Raw TCS)	.119	.139	.197	.159	.038	**.275** (.231)	.131	.196	.175	.178
CTP+GSP	.572	.574	.429	.193	.242	.276	**.578** (.564)	.379	.567	.197
ELS+MPD	.272	.265	.403	.359	.061	.293	.270	**.412** (.387)	.301	.366
Criterion 1[a]	.514	.524	.431	.223	.180	.339	.524	.390	**.533** (.517)	.235
Criterion 2[b]	.129	.128	.319	.314	.222	.245	.129	.336	.167	**.378** (.350)

Note. Rows are criteria; columns are weights corrected for range restriction; multiple $\underline{R}$ for optimal weights in bold; Rozeboom adjustments in parentheses.

[a] Criterion Factor 1 = CTP+GSP+TCS.

[b] Criterion Factor 2 = ELS+MPD+PFB.

As shown in Table 1.24, within MOS there is very little differential validity for Core Technical vs. General Soldiering Proficiency. Either set of weights works about as well. However, the same is not the case for the other four performance factors. Better prediction is always achieved by using the equation developed for each factor.

Summary of LVII Validity Estimates

In general, in spite of the small samples for each MOS and the necessity of regarding all mean criterion differences as error (i.e., standardizing criterion scores within MOS), the validities for ASVAB and ABLE were as high, or higher, for predicting second-tour performance as for predicting first-tour performance. While unit weights did not weaken the validities for ASVAB, they did constrain the predictive accuracy of ABLE.

A consistent finding from the hierarchical analysis is that for the Core Technical Proficiency, General Soldiering Proficiency, and Effort/Leadership criteria, the optimal predictor battery is never composed of only ASVAB or only ABLE factor scores. For example, the Dependability factor from the ABLE is a consistent predictor of the "can do" component of performance.

Finally, based on the above analyses, there appears to be more differential validity across MOS for the second-tour samples than was found during the analyses of the first-tour data in CVI.

All of these issues can be analyzed more rigorously when the larger samples and fuller set of predictor measures from the second-tour longitudinal (LVII) validation are analyzed.

Prediction of Second-Tour Performance From the Trial Battery and From First-Tour Performance

The fundamental research designs for Project A and Career Force include the concept of combining successive pieces of information from (a) predictor tests administered at entry, (b) measures of performance during training, and (c) measures of first-tour job performance to predict individual performance in the second tour of duty. Pending fuller analyses as larger samples become available from later data collections, a preliminary explanation was conducted with the small samples from CVII available at this stage.

These analyses of CVI and CVII data examine the relationship of ASVAB scores (from tests given at the time recruits entered the Army), the CVI predictor scores (i.e., the Project A CVI Trial Battery, the preliminary version of the Experimental Predictor Battery, given during the first tour), and first-tour job performance scores to second-tour CVII job performance scores. Two complications with these initial analyses were that available sample sizes were extremely small, and it was unclear exactly how to account for range restriction for a sample of this type.

There were 121 soldiers in Batch A MOS who had been assessed on at least a subset of measures during the CVI and CVII data collections. Not all 121 soldiers had complete CVI and CVII data; the lowest number available for a given combination of CVI and CVII measures was 102. Table 1.25 shows the number of soldiers who had CVI and CVII data, by MOS.

Table 1.25

Numbers of Soldiers With CVI and CVII Data by MOS: Initial Sample

MOS		N
11B	Infantryman	8
13B	Cannon Crewmember	26
19E	M60 Armor Crewman	4
31C	Single Channel Radio Operator	8
63B	Light-Wheel Vehicle Mechanic	25
71L	Administrative Specialist	15
88M	Motor Transport Operator	7
91A/B	Medical Specialist	15
95A	Military Police	13
	Total	121

Measures

The second-tour performance criterion CVII measures used in the exploratory analysis were the raw and residual scores for the five constructs first identified during the first-tour Concurrent Validation, and confirmed by the CVII modeling analysis. Predictor measures came from the ASVAB, from the Project A CVI Trial Battery, and from first-tour job performance measures. Because of the extremely limited sample sizes, the least-squares weights developed for the CVI criterion constructs were used rather than developing new weights for CVII criterion constructs.

Analysis and Results

CVI predictor scores were correlated with the CVII criterion scores in two ways: (a) Correlations were computed within each MOS and these values were averaged (weighted by N), and (b) correlations were computed across the total sample. Correlations with CVII criteria were computed separately for the ASVAB, spatial, computer-administered, ABLE, AVOICE, and JOB composites and for the CVI criterion scores. Correlations were also computed for the ASVAB plus each of the other predictor sets from the Trial Battery and the CVI criteria. When the CVI criteria were combined with any of the other predictor scores, they were standardized within MOS (using the larger CVI samples to compute standard scores) and summed to achieve equal weighting between ASVAB/Trial Battery and CVI criterion scores.

Because of the number of different points at which additional range restriction could occur, there are a number of different "populations" to which the CVII sample could be corrected. If the problem is to select second-tour soldiers from experienced first-tour personnel, then the set of all persons who are nearing completion of the first tour seems the most appropriate population.

The correlations of scores on the first-tour criteria with scores on second-tour criteria in the combined sample are shown in Table 1.26. The correlations are not corrected for restriction of range. The note for the table shows the mean of the diagonal correlations, which contains the correlations of the same criteria across first and second tour--that is, the correlation of Core Technical between first and second tour, and so on. This mean is an index of convergent validity for the set of criterion constructs. The note also shows the mean of the off-diagonal correlations--that is, the correlations between different criterion constructs across first and second tour. The difference between the mean diagonal and mean off-diagonal correlation can be thought of as an indicator of discriminant validity.

Table 1.26

Uncorrected Correlations Between CVI and CVII Raw Criterion Composites Computed Across MOS: Initial Sample

CVI Criterion Composite	CVII Criterion Composite				
	CTP	GSP	ELS	MPD	PFB
Core Technical Proficiency	.47	.48	.22	.10	.08
General Soldiering Proficiency	.47	.43	.36	.13	.17
Effort and Leadership	.19	.07	.30	.19	.13
Maintaining Personal Discipline	.06	.14	.16	.26	.19
Physical Fitness and Military Bearing	.00	-.04	.15	.15	.48

Note. Ns = 102-121. Mean diagonal value = .39; mean off-diagonal value = .17.

The correlations of predicted scores based on CVI weights for ASVAB and Trial Battery composites and CVI criterion scores with CVII criteria are shown in Table 1.27. On the whole, of all the predictors, the CVI criterion scores have the highest correlations with CVII criterion scores. However, adding the ASVAB and the ASVAB plus Trial Battery composite scores to CVI scores does increment the CVI validity coefficients.

The ASVAB validities follow the familiar pattern of predicting the two "can do" criteria, but not predicting the "will do" criteria very well. The JOB unexpectedly did the best job of predicting Maintaining Personal Discipline.

In sum, the results with this small initial sample provide evidence that ASVAB scores, weighted on the basis of regression estimates for predicting first-tour performance, predict second-tour "can do" performance with substantial validity. The results also provide evidence of convergent and

Table 1.27

Correlations Between CVI Weighted Predictor Composites, CVI Criterion Composites, and CVII Criterion Composites for Raw Scores, Computed Across MOS: Initial Sample

Predictor and CVI Criterion Composites and Combinations	CVII Criterion Composite				
	CTP	GSP	ELS	MPD	PFB
ASVAB	.33	.42	.11	-.05	.11
CVI Performance	.47	.43	.30	.26	.48
ASVAB+CVI Performance	.51	.51	.33	.26	.47
Computer Tests	.23	.13	-.01	-.04	.10
ASVAB+Computer Tests	.37	.41	.13	.05	.12
ASVAB+Comp. Tests+CVI Performance	.52	.51	.33	.27	.46
AVOICE	.15	.16	.06	-.02	.06
ASVAB+AVOICE	.43	.44	.14	.00	.13
ASVAB+AVOICE+CVI Performance	.54	.52	.33	.27	.46
JOB	.12	.00	.19	.30	.12
ASVAB+JOB	.33	.41	.16	.20	.16
ASVAB+JOB+CVI Performance	.51	.51	.34	.31	.48
Spatial	.47	.41	.14	-.01	.04
ASVAB+Spatial	.41	.43	.10	-.06	.11
ASVAB+Spatial+CVI Performance	.52	.51	.33	.26	.46
ABLE	.10	.01	.21	.15	.29
ASVAB+ABLE	.34	.41	.22	.12	.25
ASVAB+ABLE+CVI Performance	.51	.52	.36	.30	.47

Note. Ns = 102-121. Correlations are uncorrected for range restriction. Coefficients do not require shrinkage adjustments. CVI criterion scores and predictor composites were summed.

discriminant validity of the first-tour job performance for predicting second-tour job performance criteria.

Future analyses of the LVI Experimental Predictor Battery and LVII criterion scores will provide better indications of the new predictors' relationships with second-tour performance.

SUMMARY OF PROJECT EFFORTS FOR YEAR THREE

As described in the third annual report (Campbell & Zook, 1994b), the Project had four main objectives during FY 92:

(1) Complete the Longitudinal Second-Tour (LVII) data collection.

(2) Prepare the LVII data file for analysis.

(3) Analyze the LVII criterion data to develop the basic performance scores.

(4) Model the covariance structure of the LVII performance scores.

In general, because the LVII results could be considered as a major replication of CVII, a great deal of effort was devoted to using the LVII sample data in a confirmatory way. That is, when possible, the CVII results were used as a hypothesis to be tested in LVII.

Longitudinal Validation Second-Tour Data Collection

The LVII data collection administered second-tour criterion measures to soldiers in the longitudinal validation sample who had reenlisted for a second tour and who were available for assessment at a specified set of data collection locations. Although this data collection involved substantially fewer soldiers than the CVI or LVI data collections, it posed a number of special challenges. Having to locate and test particular individual soldiers, especially when there were relatively few to begin with, was a difficult task. However, despite a major deployment of U.S troops to Southwest Asia (Operation Desert Shield/Storm), LVII data were collected from 1,577 soldiers.

A list of the instruments administered in the LVII data collection is provided in Table 1.28. Most of the instruments served as second-tour performance criterion measures; several other instruments (e.g., the Background Information Form) provided supplemental data for the project.

The original project plan called for the LVII data collection to take place July-December 1991. Second-tour criterion data were to be collected from at least 150 soldiers in each of nine MOS (the Batch A group designated in previous data collections).

Even before the deployment of troops to Southwest Asia, the anticipated data collection problems included difficulty projecting future location of soldiers targeted for testing because of frequent reassignments, and limited access due to training or alert status, leave, and so forth. The problems were compounded by a tasking system which requires that Troop Support Requests (TSRs) be submitted well in advance of data collection (135 days up to 500 days for U.S. Forces Command, less for the U.S. Training and Doctrine Command). Moreover, before detailed data collection planning activities began, the Army was starting to respond to directives to downsize and to reduce the proportion of troops stationed in Germany.

These concerns led to the following strategy for maximizing the number of LVII subjects. In May 1990, analyses were conducted to determine the number of Project A soldiers who were still in the Army and their locations.

Table 1.28

LVII Data Collection Instruments

Performance Criterion Instruments

- Job Knowledge Tests
- Hands-On Tests
- Performance Rating Scales (completed by supervisors)

  - Army-Wide Booklet
  - MOS-Specific Booklet
  - Combat Performance Prediction Scales
  - Combat Performance Questionnaire (Operation Desert Shield/Storm), administered if applicable

- Personnel File Form (PFF)
- Situational Judgment Test (SJT)
- Supervisory Simulation Exercises

  - Personal Counseling
  - Disciplinary Counseling
  - Training

Supplemental Instruments

- Background Information Form
- MOS-Specific Job History Questionnaire
- Supervisory Experience Questionnaire
- Army Job Satisfaction Questionnaire (AJSQ)
- Assessment of Background and Life Experiences (ABLE)
- Leader and Unit Attitudes Questionnaire

It was clear that sample size requirements would not be met if only soldiers having predictor and first-tour criterion data were tested. Accordingly, the decision was made to test soldiers for whom predictor and/or first-tour criterion data were available. Only soldiers with Project A data were eligible. An additional consideration was the fact that, shortly after the LVI data collection ended, MOS 31C began declining in strength because certain radio equipment was being phased out. The collection of hands-on data is inordinately resource-intensive for small numbers of examinees; consequently, hands-on tests were dropped from the performance measures for the 31C soldiers.

The May 1990 analyses also indicated that appreciable concentrations of Project A soldiers were stationed in locations other than those identified in the original research plan. Accordingly, requests for troop support were written to include some of these new sites. Then, lists of all soldiers eligible for testing were electronically matched with each installation's own personnel files, to obtain the most accurate identification of soldiers qualified for testing at each location.

Data Collection Schedule

The original research plan, calling for LVII data collection July-December 1991, was adjusted to accommodate the interests of supporting commands. It was agreed that test sites could be scheduled to conduct testing as early as May 1991 and as late as February 1992.

This data collection strategy had been established before hostilities involving U.S. troops in Southwest Asia arose. The U.S. Forces Command, which was tasked to provide the majority of LVII soldiers, invoked a moratorium on research support in September 1990. The moratorium was lifted in April 1991, and the data collection schedule was again modified. The final schedule is shown in Table 1.29. Four data collection teams were sent to Germany, whereas one team of data collectors was sent to each of the other test sites. The first LVII data collection occurred in June 1991 and the last in July 1992.

Table 1.29

LVII Data Collection Schedule

Command	Location	Test Dates
		1991
USAREUR	Germany	7 June - 27 June
USAREUR	Germany	5 July - 2 August
USAREUR	Germany	5 July - 3 August
Eighth Army	Republic of South Korea	5 July - 9 August
USAREUR	Germany	September - October
HSC	Fort Sam Houston, TX	October
FORSCOM	Fort Lewis, WA	9 December - 19 December
		1992
FORSCOM	Fort Drum, NY	13 January - 24 January
TRADOC	Fort Bliss, TX	20 January - 31 January
MDW & AMC	Fort Belvoir, VA	February
TRADOC	Fort Knox, KY	2 March - 6 March
FORSCOM	Fort Bragg, GA	16 March - 3 April
TRADOC	Fort Benning, GA	31 March - 3 April
FORSCOM	Fort Riley, KS	6 April - 10 April
FORSCOM	Fort Hood, TX	4 May - 15 May
FORSCOM	Fort Campbell, KY	11 May - 15 May
FORSCOM	Fort Carson, CO	1 June - 5 June
FORSCOM	Fort Stewart, GA	15 June - 23 June
TRADOC	Fort Polk, LA	13 July - 16 July

USAREUR U.S. Army Europe
HSC Health Services Command
FORSCOM Forces Command
TRADOC Training and Doctrine Command
MDW Military District of Washington
AMC Army Materiel Command

Composition of the teams, in terms of project staff, varied from location to location. Generally however, each test site was staffed with a team comprised of the following personnel:

1	Test Site Manager (TSM)
1-2	Hands-on Managers (HOMs)
3	Test Administrators (TAs)

All of these positions were filled by permanent employees of the contractor consortium; often a representative of the Army Research Institute was present during the testing. The Army installations also provided personnel to help support the data collection activities. In addition to the test site POC, each test site provided eight senior NCOs for each MOS (except 31C) to administer and score the hands-on tests, and two to four NCOs to fill general supporting roles (e.g., to track down soldiers who failed to report for testing and to handle problems with defective equipment).

Data Collection Team Training

One day of classroom training and considerable follow-up on-the-job training were provided to TAs for the written test and supervisor rating procedures. One to two days of additional training were provided to each TA for each subordinate role a TA was responsible for playing in the Situational Judgment Test. Two documents were developed to support TA training: the Test Administrator's Manual and the Supervisory Role-Play Exercises Administration Manual.

NCO hands-on scorers were trained the day before the administration of the hands-on tests to soldiers in a given MOS. The training followed the same basic procedures as those that had been used in the CV and LVI/CVII data collections (R. Campbell, 1985).

Various procedures and documents were used to handle completed data collection instruments before they were shipped to the facility where they would be processed and keypunched. Test site personnel checked measures for completeness and legibility, and documented explanations for data that were incomplete or otherwise anomalous. Transmittal documents were used to help ensure that data could be tracked once it left the test site.

After data collection at a given location was completed, the TSM prepared and submitted a report to ARI.

Development of Basic Scores for LVII Performance Measures

The LVII performance criterion measures have been described in detail elsewhere (Campbell, 1989; Campbell & Zook, 1990). They were originally administered to second-tour soldiers in the CVII sample and were subsequently revised in preparation for administration to the LVII sample.

Analyses of the data from the LVII sample had three major objectives: (a) examine and evaluate the psychometric properties of the LVII measures, (b) compare the psychometric properties of the LVII scores with the CVII scores, and (c) develop the basic criterion scores to be used in modeling second-tour performance.

Job Knowledge and Hands-On Tests

In earlier research a set of 28-30 tasks had been selected for performance measurement in each MOS. All tasks were assessed using a written job knowledge test format. Performance on a subset (14-17) of the tasks was assessed using a hands-on performance test format.

The full set of tasks included (a) common tasks, basic soldiering tasks that all soldiers are expected to know how to perform (e.g., first aid, personal weapons, map reading), and (b) MOS-specific tasks, central to the jobs of the soldiers working in a given MOS and typically unique to that MOS.

Some tasks are performed differently depending upon the type of equipment a soldier uses (e.g., an M16A1 rifle versus an M16A2 rifle). To deal appropriately with such situations, tracked (i.e., parallel) tests were prepared for tasks where the equipment varied.

Before the CVII measures could be used again, technical currency reviews were conducted. Revisions were made to test items and to supporting graphics and handouts as necessary.

Scoring Adjustments. Specifications for the basic scores for the LVII job knowledge and hands-on measures depended largely on previous work in CVI, CVII, and LVI. Job knowledge and hands-on task scores were calculated as percent-correct (or percent-GO) scores at all score levels. The data for tracked tests were examined for evidence of level and dispersion differences between tracks and no anomalous differences were found.

As with the previous data collections, hands-on test scores were standardized by site at the task level to control for site differences. One adjustment affected only the job knowledge tests; that is, between one and four items per MOS were dropped because of doctrinal changes subsequent to CVII.

Table 1.30 shows the overall number of items in the job knowledge component for each MOS and the range of items per task test. Table 1.31 shows the overall number of steps in the hands-on component for each MOS and the range of steps per task test.

After data editing, four levels of scores (Tasks, Functional Categories, Task Factors, and Task Constructs) were constructed. They are as depicted in Figure 1.10 in an earlier section of this chapter.

At each level of aggregation, hierarchical scores were computed using task-level data. That is, each category, factor, and construct score was computed by calculating the mean percentage of items correct (or percentage of steps passed) across all constituent tasks.

Final Basic Scores for Job Knowledge and Hands-On Measures. The descriptive statistics calculated across MOS for both the Task Construct and Task Factor scores do not differ much from the results for the CVII soldiers tested (reported in Campbell & Zook, 1990).

Table 1.30

Number of LVII Job Knowledge Tasks and Items by MOS

MOS		No. of Tasks	Items Dropped	Total Items	Items Per Task	Average Items Per Task
11B	Infantryman[a]	29	2	128	2-12	4.4
13B	Cannon Crewmember[a]	30	3	119-120	2-8	4.0
19K	M1 Armor Crewman	28	4	142	3-12	5.1
31C	Single Channel Radio Operator[a]	30	1	111-112	3-5	3.7
63B	Light Wheel Vehicle Mechanic	27	2	102	2-6	3.8
71L	Administrative Specialist[a]	30	2	125	2-12	4.2
88M	Motor Transport Operator	30	1	119	3-12	4.0
91A/B	Medical Specialist	30	3	113	2-6	3.6
95B	Military Police	29	4	109	2-7	3.8

[a] One or more task tests were tracked; tracked tests do not necessarily have the same number of items.

Table 1.31

Number of LVII Hands-On Tasks and Steps by MOS

MOS		No. of Tasks	Total Steps	Steps Per Task	Average Steps Per Task
11B	Infantryman	9	121	5-31	13.4
13B	Cannon Crewmember[a]	12	258-259	7-67	21.5-21.6
19K	M1 Armor Crewman	10	167	8-37	16.7
63B	Light Wheel Vehicle Mechanic[a]	8	142	7-44	17.8
71L	Administrative Specialist[b]	14	140-146	2-44	10.0-10.4
88M	Motor Transport Operator[a]	10	193-195	4-44	19.3-19.5
91A/B	Medical Specialist[a]	13	216	6-44	16.6
95B	Military Police[a]	10	223-227	7-37	22.3-22.7

[a] One or more task tests were tracked; tracked tests do not necessarily have the same number of steps.
[b] One task was scored on a continuous scale; it is not included in calculating total steps, steps per task, or average steps per task.

Task Factor (otherwise known as CVBITS) scores had been used in the performance modeling exercises conducted for CVI and LVI; however, Task Construct scores (i.e., MOS-Specific and General) were used for this purpose in CVII. Although Task Factors preserve somewhat more information than the more highly aggregated Task Construct scores, they have the disadvantage of differing across MOS as to the availability of each of the six scores. This

problem is compounded by the considerably smaller sample sizes available for the two second-tour data collections relative to the two first-tour data collections. Moreover, in both CVI and LVI, the Technical Task Factor score invariably loaded on the Core Technical Proficiency performance construct while the other five Task Factor scores invariably loaded on the General Soldiering Proficiency performance construct. Therefore, the two Task Construct scores were selected for use in the LVII performance modeling exercise.

Performance Rating Scales

As reported previously (Campbell, 1989), the second-tour performance rating scales (with the exception of the Combat Performance Questionnaire) were developed by revising and adapting rating scales developed for first-tour soldiers. Based on results of the CVII data analyses, additional minor modifications were made to these three sets of scales: the Army-Wide ratings, the MOS-Specific ratings, and the Combat Performance Prediction scales.

Army-Wide Ratings. The Army-Wide rating booklet included 12 behavior-based dimensions, seven task-based leadership dimensions, a rating of overall effectiveness, and a rating of senior NCO potential.

Raters in the CVII sample had tended to make frequent use of the highest rating scale values, suggesting that the scale behavioral anchors may have been too lenient for more experienced soldiers. In the LVII sample, the behavioral anchors for most rating dimensions were revised to make the scale values reflect a slightly higher level of performance.

Mos-Specific Ratings. The MOS-Specific rating booklets included from 7 to 14 technically oriented behavior-based dimensions and a rating of overall MOS effectiveness. A set of scales suitable for second-tour MOS 19K soldiers were developed by adapting the second-tour MOS 19E scales that had been used in CVII. In five of the nine MOS, one or two of the MOS-specific dimensions measured some aspect of leadership (e.g., Leading the Team for MOS 11B). As with the Army-wide rating dimensions, the CVII behavioral anchors for most MOS-specific rating dimensions were revised to reflect slightly higher levels of performance.

Combat Performance Prediction Scales. The Combat Performance Prediction scales consisted of 14 items which depict examples of soldier behaviors under combat conditions. The rater's task was to estimate the likelihood that the ratee would behave as described in the behavioral example. Ratings were made on a 7-point scale ranging from very likely to very unlikely. The items were a subset of the 40 items that appeared on the original CVI version of the Combat Performance Prediction scales. Unlike the LVI/CVII data collections, LVII Combat Performance Prediction scale ratings were collected for both male and female soldiers.

Summary of Ratings Data. Across all nine MOS, two or more ratings were obtained for 75 percent of the soldiers (1,194 of 1,595) and at least one rating was obtained for 94 percent of the sample (1,494 of 1,595). The soldiers who received ratings averaged 1.82 raters per ratee.

Substantive analyses for the Army-wide and Combat Performance Prediction scale ratings were carried out on the total sample; MOS-specific

ratings were, of course, analyzed separately by MOS. The analyses for the Army-wide and MOS-specific rating scales focused first on the distributions of the individual ratings and reliability estimates. This was followed by principal factor analyses with varimax rotation to determine the composition of basic scores.

Overall, the LVII rating distributions were as expected. The means were generally lower than in CVII and the variability was similar. The interrater reliability for the LVII ratings was almost exactly the same as that found in the LVI and CVII research.

Several factor analyses were conducted on the LVII sample. Army-wide ratings on the nine second-tour nonleadership dimensions were intercorrelated and factor analyzed so that the LVI and LVII factor structures could be compared. The same procedure was followed for all 19 of the Army-wide dimensions. For this analysis, the factor structure obtained in the LVII sample was compared to the factor structure obtained in the CVII sample.

The similarity of the rotated factor structures for LVI vs. LVII for the nine nonleadership/supervision dimensions that are common to the first-tour and second-tour rating scales was striking.

The four-factor rotated solutions obtained in the LVII and CVII samples for all 19 scales are shown in Table 1.32. Both include three factors that are quite similar to the three LVI factors, plus a separate leadership/supervision factor.

Basic Scores. Factor analyses of the Army-wide ratings suggest that the four-factor CVII solution is appropriate for LVII. Accordingly, the four composites shown in Table 1.33 and the overall effectiveness rating were used as basic scores for the LVII Army-wide rating data.

Because the factor analysis results did not indicate multiple factors for any of the MOS-specific rating analyses, a unit-weighted composite of all dimension ratings for each MOS was used as the basic score. This is identical to the scoring system used in CVII, and yielded comparable reliability estimates.

Results of the principal component analysis for the Combat Performance Prediction scales on the combined LVII sample confirmed the findings that were obtained in LVI and CVII. Specifically, two factors were identified; however, the second factor was simply a reflection of the first (i.e., it was comprised of the negatively worded items). The 14 items were summed to form a single basic score.

The mean LVII Combat Performance Prediction scale composite score and standard deviation were virtually identical to the mean and standard deviation of the supervisor ratings of soldiers in the CVII sample.

Table 1.32

Comparison of LVII and CVII Army-Wide Factor Analysis[a] Results: All Dimensions

	Factor Loadings (LVII/CVII)			
Dimension	1	2	3	4
1. Technical Knowledge/Skill	.47/.41	.23/.24	.26/.22	.56/.65
2. Effort	.45/.39	.34/.31	.26/.27	.58/.68
3. Supervising	.63/.57	.22/.21	.24/.28	.42/.53
4. Following Regs/Orders	.32/.29	.63/.63	.29/.30	.31/.36
5. Integrity	.38/.32	.58/.66	.24/.22	.34/.37
6. Training/Development	.60/.52	.20/.24	.27/.27	.38/.52
7. Maintain Equipment	.36/.32	.27/.33	.32/.25	.38/.50
8. Physical Fitness	.17/.20	.14/.18	.53/.60	.16/.19
9. Self-Development	.48/.41	.29/.27	.41/.44	.32/.48
10. Consideration for Subord	.61/.47	.40/.44	.16/.26	.28/.40
11. Military Bearing	.26/.30	.32/.34	.62/.63	.12/.22
12. Self-Control	.16/.17	.57/.56	.20/.18	.07/.09
13. Role Model	.51/.53	.37/.40	.56/.51	.25/.31
14. Communication	.61/.62	.39/.34	.22/.23	.26/.35
15. Personal Counseling	.74/.72	.24/.31	.27/.26	.11/.19
16. Monitoring	.68/.63	.18/.31	.30/.22	.30/.41
17. Organizing Missions/Operations	.66/.70	.22/.26	.27/.20	.30/.36
18. Personnel Administration	.65/.63	.28/.20	.22/.24	.17/.29
19. Performance Counseling	.72/.72	.22/.20	.23/.29	.24/.32
Percent Common Variance	45.3/37.6	25.4/20.3	18.2/16.9	16.9/25.3

Note. Sample size is 1,388 for LVII and 823 for CVII. CVII analyses based on supervisor ratings only.

[a] Principal factor analysis, varimax rotation.

Table 1.33

Composition of LVII Army-Wide Rating Composites

Percent Common Variance Accounted for by Relevant Factor	Composite Name	Dimensions Included
45.3	1. Leading/Supervising	Supervising Training/Development Consideration for Subord Communication Personal Counseling Monitoring Organizing Missions/Opers Personnel Administration Performance Counseling
25.4	2. Personal Discipline	Following Regs/Orders Integrity Self-Control
16.9	3. Technical Skill/Effort	Technical Knowledge/Skill Effort Maintain Equipment
18.2	4. Physical Fitness/ Military Bearing	Military Bearing Physical Fitness

Note. Two dimensions were not included in any composites: Acting as a Role Model and Self-Development.

Administrative Measures: The Personnel File Form

The LVII Personnel File Form was used to gather self-reports of archival/administrative information dealing with personnel actions reflective of individual performance. The first-tour versions (CVI and LVI) of the PFF requested information regarding (a) evidence of exemplary performance, including awards and memoranda/certificates of appreciation, commendation, and achievement; (b) receipt of disciplinary actions (i.e., Articles 15 and flag actions); and (c) test results, including Physical Readiness test scores, individual weapon qualification scores, and Skill Qualification Test scores.

The original second-tour version of the PFF developed for CVII included these same types of variables and added others. The additional items were related to education (military training and civilian college courses) and

promotions (e.g., how often recommended for accelerated promotion, number of promotion board points received). Another modification was to distinguish between awards, memoranda, and disciplinary actions received while in grades E-1 through E-3 and those received while in grades E-4 and above.

Before being administered to the LVII sample, the second-tour PFF was revised in several minor ways. Most of these revisions were intended to increase the interpretability/accuracy of responses and to reduce the amount of missing data. For example, the PFF response format was changed so that soldiers could indicate if they had earned more than one Army Achievement Medal.

Descriptive Statistics and Intercorrelations. Means and standard deviations for the administrative indices of performance are presented in Table 1.34. The corresponding descriptive statistics for CVII are not comparable for the Awards and Certificates score because of response format differences between the CVII and LVII instruments. Otherwise, the means and standard deviations for the LVII and CVII scores are very similar.

Subsequent analyses suggested that the basic scores tentatively derived for the PFF satisfactorily captured the useful information on that form. Consequently, they were made available for use in the second-tour performance modeling exercise.

Table 1.34

Administrative Indices Descriptive Statistics for LVII and CVII

Measure		N	Mean	SD	Range
Awards and Certificates[a]	CVII	928	10.53	5.63	0-44
	LVII	1,577	14.69	6.79	0-40
Disciplinary Actions	CVII	930	.42	.87	0-8
	LVII	1,577	.37	.76	0-6
Physical Readiness Score	CVII	998	250.11	30.68	121-300
	LVII	1,522	248.81	31.27	23-288
Weapon Qualification	CVII	1036	2.52	.67	1-3
	LVII	1,565	2.58	.67	1-3
Promotion Rate	LVII	1,513	100.00	7.79	61-128
Promotion Rate	CVII	901	100.14	8.09	67-121
(CVII Scoring)	LVII	1,513	99.98	7.48	57-121

[a] Differences in LVII and CVII results reflect differences in PFF response format.

Situational Judgment Test (SJT)

The SJT was designed to evaluate the effectiveness of NCO judgments concerning what the NCO should do in difficult supervisory situations. Thus, the SJT can be viewed as a job knowledge test pertaining to the leadership/ supervision components of second-tour positions.

As reported previously (Campbell, 1989), development of the SJT involved asking groups of soldiers similar to the target soldiers (i.e., at the level of SP4/SP5/Sergeant) to describe a large number of difficult but realistic situations that Army first-line supervisors face on their jobs. After a large number of these situations had been generated, a wide variety of possible actions (i.e., response alternatives) for each situation were gathered, and ratings of the effectiveness of each of these actions were collected from both experts (senior NCOs) and the target group. These effectiveness ratings were used to select situations and response alternatives to be included on the SJT.

The initial version of the SJT, which was administered to the CVII sample, consisted of 35 items. Because the CVII data analysis results indicated that the SJT was a promising measure of supervisory performance, this test was lengthened to 49 items for the LVII data collection to increase the internal consistency reliability and make it easier to identify SJT subscores.

Because the SJT had proven to be relatively difficult for the CVII sample, an effort was made to select relatively less difficult additional items to include in the lengthened test for the LVII administration.

Analysis Procedure. Procedures for scoring the SJT were identical to those used in CVII. Five different scores were computed. The most straightforward was a simple number correct score. The second score involved weighting each response alternative by the mean effectiveness rating given to that response alternative by the expert group.

Two analogous scoring procedures based on respondents' choices for the least effective response to each situation were also used.

The fifth score involved combining the choices for the most and the least effective response alternative into one overall score. For each item, the mean effectiveness of the response alternative each soldier chose as the least effective was subtracted from the mean effectiveness of the response alternative they chose as the most effective. Because it is actually better to indicate that less effective response alternatives are the least effective, this score can be seen as not a "difference" score but a simple sum.

Each of these scores was computed twice for the LVII soldiers, once using all 49 SJT items and once including only the 35 SJT items that had been administered to the CVII sample as well.

Descriptive statistics and internal consistency reliabilities were computed for each of the five scoring procedures for both the 49-item and the 35-item versions of the SJT. Intercorrelations were computed among the five scores generated by the five different scoring procedures for the 49-item SJT only. Finally, item analyses were conducted for each of the scoring procedures. (See Campbell & Zook, 1994b.)

Development of Subscales. Efforts to identify distinct SJT subscores in the CVII sample had not been particularly successful. However, the LVII version of the SJT contained almost 40 percent more items, and it was possible that a more interpretable solution would be found using the LVII data. In addition, a content analysis of the SJT items conducted by Hanson and Borman (in press) revealed some promising new subscales. Consequently, the dimensionality of the SJT for the LVII sample was investigated both rationally and empirically, with the primary goal to develop a set of more homogeneous SJT subscores.

The item-level responses from the LVII sample were well distributed across the response alternatives for each item, which suggests that the correct responses to SJT items were not obvious. All of the scoring procedures resulted in a reasonable amount of variability in both the LVII and CVII samples. The most reliable score for both samples is M-L Effectiveness, probably because this score contains the most information.

For the 49-item SJT, in which the maximum possible M-Correct score is 49, the mean in the LVII sample is only 25.84, indicating that this longer version of the SJT was also relatively difficult.

Based on the full array of descriptive statistics, the M-L Effectiveness score appeared to provide the best summary of the information contained in the SJT responses. Thus, all remaining analyses focused on the M-L Effectiveness score, which became the SJT Total Score.

The rational/empirical analysis of the item covariances resulted in six factor-based subscales that contained between six and nine items each. Six remaining items were not included in any subscale. Definitions of these factor-based subscales are presented in Table 1.35. These subscales have potential for more clearly delineating the leadership/supervision aspects of the second-tour soldier job. They are included in one of the major alternative models of second-tour performance to be evaluated in subsequent confirmatory analyses.

Supervisory Simulation Exercises

The supervisory simulation measures were designed to assess areas of second-tour job performance that deal with specific components of supervisor/subordinate interaction. These areas included personal counseling, disciplinary counseling, and one-on-one training. A trained evaluator (role-player) acted out the role of a subordinate to be counseled or trained and the examinee assumed the role of a first-line supervisor who was to conduct the counseling or training. In each exercise, evaluators scored the examinees on a number of rating scales.

The subordinate and supervisor roles were essentially the same as those used in the CVII data collection. The role-players who assumed the role of the subordinate in each of these exercises were trained to play the roles in a standardized fashion.

The rating system used to evaluate LVII examinees was modified in several ways from CVII. First, the CVII analyses identified the scales that (a) appeared to be difficult to rate, (b) conceptually redundant, and/or (c) not correlated with other rated behaviors in meaningful ways. These behavior

Table 1.35

Situational Judgment Test: Definitions of Factor-Based Subscales

1. Discipline soldiers when necessary (Discipline). This subscale is made up of items on which the most effective responses involve disciplining soldiers, sometimes severely, and the less effective responses involve either less severe discipline or no discipline at all. (Six items.)

2. Focus on the positive (Positive). This subscale is made up of items on which the more effective responses involve focusing on the positive aspects of a problem situation (e.g., a soldier's past good performance, appreciation for a soldier's extra effort, the benefits the Army has to offer). (Six items.)

3. Search for underlying reasons (Search). This subscale is made up of items on which the more effective responses involve searching for the underlying causes of soldiers' performance or personal problems rather than reacting to the problems themselves. (Eight items.)

4. Work within the chain of command and with supervisor appropriately (Chain/Command). For a few items on this subscale the less effective responses involve promising soldiers rewards that are beyond a direct supervisor's control (e.g., "comp" time). The remaining items involve working through the chain of command appropriately. (Six items.)

5. Show support/concern for subordinates and avoid inappropriate discipline (Support). This subscale is made up of items where the more effective alternatives involve helping the soldiers with work-related or personal problems and the less effective responses involve not providing needed support or using inappropriately harsh discipline. (Eight items.)

6. Take immediate/direct action (Action). This subscale is composed of items where the more effective response alternatives involve taking immediate and direct action to solve problems and the less effective response alternatives involve not taking action (e.g., taking a "wait and see" approach) or taking actions that are not directly targeted at the problem at hand. (Nine items.)

---

ratings were dropped and anchors for some of the remaining scales were changed. The scale itself was expanded from 3 to 5 points. The overall effectiveness rating was retained, but the overall affect and fairness rating scales were eliminated. Thus, examinees were rated on each exercise on from 7 to 11 behavioral scales and on one overall effectiveness scale.

Another important difference between the CVII and LVII measures was the background of the evaluators. The smaller size of the LVII data collection allowed for the selection and training of role-players/evaluators who had been formally educated as personnel researchers and who were employed full-time by organizations in the project consortium. In contrast, the scope of the LVI/CVII data collection had required the hiring of a number of temporary employees to serve as role-players.

Descriptive analyses were conducted, followed by a series of factor analyses. Maximum likelihood factor analyses with oblique rotations were performed within each exercise. The purpose of the factor analyses was to identify the content of basic criterion scores for each of the simulation exercises.

The median and the range of the scale means and the median and the range of the scale standard deviations, for each exercise, indicated that (a) there is a reasonable amount of variation in the scale ratings, (b) none of the scale ratings show a floor effect, and (c) a reasonable number of the ratings do not show a ceiling effect.

Personal Counseling Exercise. Table 1.36 presents the pattern of matrices resulting from the factor analyses of the standardized and raw score Personal Counseling exercise ratings that specified two factors. The two-factor structure was preferred over the one- or three- (or more) factor structures based on the superior simple structure and interpretability of the rotated two-factor pattern matrix. Factor 1 was labeled "Communication/ Interpersonal Skill," and Factor 2 was labeled "Diagnosis/Prescription."

As indicated by the notations in Table 1.36, the factor analysis results for LVII did not exhibit the same pattern as that obtained in CVII.

Disciplinary Counseling Exercise. Table 1.37 presents the pattern of matrices resulting from the factor analyses of the standardized and raw scale Disciplinary Counseling exercise ratings that specified three factors. The three-factor structure was preferred over the one-, two-, or four- (or more) factor structures based on the superior simple structure and interpretability of the rotated three-factor pattern matrix. Factor 1 was labeled "Structure," Factor 2 was labeled "Interpersonal Skill," and Factor 3 was labeled "Communication."

Training Exercise. Table 1.38 presents the pattern matrices resulting from the factor analyses of the standardized and raw scale Training exercise ratings that specified two factors, which were labeled "Structure" and "Motivation Maintenance."

Basic Scores. Scales were assigned to composite scores based on the patterns of their relative factor loadings in the factor structure for each exercise. This procedure resulted in empirically derived basic scores for each exercise that seemed to have considerable substantive meaning.

Across all exercises, each basic composite score was generated by (a) standardizing the ratings on each scale within each evaluator, (b) scaling each standardized rating by its raw score mean and standard deviation, and (c) calculating the mean of the transformed (i.e., standardized and scaled) ratings that were assigned to that particular basic criterion composite. The ratings were standardized within evaluator because (a) each evaluator rated examinees in only some MOS and (b) there was more variance in mean ratings across evaluators than there was in mean ratings across MOS. The standardized ratings were scaled with their original overall means and standard deviations so that each scale would retain its relative central tendency and variability.

Table 1.36

LVII Personal Counseling Exercise Scales and Factor Analysis Results[a]

Scale	Factor 1	Factor 2	h^2
Communication/Interpersonal Skill			
1. States the purpose of the counseling session clearly and concisely.[1]	.45	-.04	.18
2. Gives the subordinate positive feedback for his/her overall good past performance.[1]	.74	-.10	.48
3. Explains what the solider did wrong and why it was or can be a problem.[1]	.38	-.06	.12
7. Maintains eye contact during the interview.[2]	.30	.14	.16
8. Behaves in a manner that demonstrates support and concern for subordinate.[OMIT]	.52	.30	.54
9. Conducts the counseling session in a professional manner.[2]	.47	.12	.29
10. Maintains open communication.[2]	.13	.45	.27
Diagnosis/Prescription			
4. Asks open-ended, fact -finding questions that uncover important and relevant information.[1]	.01	.78	.61
5. Provides advice to the subordinate concerning actions that should be taken to solve problems.[1]	-.04	.87	.73
6. Sets a time or date to follow-up with the subordinate.[1]	.01	.52	.27
Omitted Item			
11. Does not interrupt the subordinate when he/she is talking.[2]	.08	.17	.05
Eigenvalue[b]	6.73	1.39	

Note. The underline indicates which composite the scale was assigned to for the construction of simulation exercise basic scores; h^2 = Communality.

[a] Maximum likelihood factor analysis with an oblique rotation. From analysis of standardized scale ratings.
[b] Eigenvalues of the first two unrotated factors.

[1] A similar (or the same) scale was assigned to the Personal Counseling - Content composite score in CVII.
[2] A similar (or the same) scale was assigned to the Personal Counseling - Process composite score in CVII.
[OMIT] A similar scale was not assigned to a composite score in the CVII analyses.

Table 1.37

LVII Disciplinary Counseling Exercise Scales and Factor Analysis Results[a]

Scale	Factor 1	Factor 2	Factor 3	h^2
Structure				
1. Remains focused on the immediate problems (i.e., the subordinate's absences and/or lying).[1]	.38	.12	-.08	.17
2. Determines an appropriate corrective action.[1]	.57	.04	-.02	.33
3. States the exact provisions of the punishment.[1]	.57	-.01	.07	.34
Interpersonal Skill				
6. Conducts the counseling session in a professional manner.[2]	.07	.72	-.02	.53
7. Defuses rather than escalates potential arguments.[2]	-.04	.67	-.02	.44
Communication				
4. Explains the ramifications of the soldier's actions.[OMIT]	.01	-.03	.82	.66
5. Allows the subordinate to present his/her view of the situation.[2]	.14	.08	.29	.14
Eigenvalue[b]	2.62	1.52	1.02	

Note. The underline indicates which composite the scale was assigned to for the construction of simulation exercise basic scores; h^2 = Communality.

[a] Maximum likelihood factor analysis with an oblique rotation. From analysis of standardized scale ratings.
[b] Eigenvalues of the first three unrotated factors.

[1] A similar (or the same) scale was assigned to the Disciplinary Counseling - Content score in CVII.
[2] A similar (or the same) scale was assigned to the Disciplinary Counseling - Interpersonal Skills score in CVII.
[OMIT] A similar item was not assigned to a composite score in the CVII analyses.

Table 1.38

LVII Training Exercise Scales and Factor Analysis Results[a]

Scale	Factor 1	Factor 2	h^2
Structure			
2. Organizes and presents the training steps in a logical sequence.	.64	-.03	.39
3. Demonstrates the task steps for the trainee.	.58	.07	.39
4. Identifies and corrects the trainee's errors.	.74	-.16	.41
5. Makes the trainee practice each movement required to perform the task.	.66	-.03	.41
6. Provides specific feedback to the trainee following good performance.	.70	.04	.53
Motivation Maintenance			
7. Provides positive feedback to the trainee following good performance.	-.01	.81	.65
8. Encourages the trainee when mistakes are made.	-.07	.80	.57
Omitted Items			
1. Presents an overview of what will be learned.	.18	.21	.13
9. Speaks in a clear, distinct, and understandable manner.	.28	.26	.25
Eigenvalues[b]	6.12	1.32	

Note. The underline indicates which composite the scale was assigned to for the construction of simulation exercise basic scores. In the CVII analyses scales similar (or identical) to those above were assigned to a single Training Exercise composite score. h^2 = Communality.

[a] Maximum likelihood factor analysis with an oblique rotation. From analysis of standardized scale ratings

[b] Eigenvalues of the first two unrotated factors.

Summary of Basic Criterion Scores

The analyses of the LVII performance measures resulted in an array of basic criterion scores which were available for the performance modeling analyses. These scores are summarized in Figure 1.12.

Hands-On Performance Test

1. MOS-specific task performance score
2. General (common) task performance score

Job Knowledge Test

3. MOS-specific task knowledge score
4. General (common) task knowledge score

Army-Wide Rating Scales

5. Overall effectiveness rating
6. Leadership/supervision composite
7. Technical skill and effort composite
8. Personal discipline composite
9. Physical fitness/military bearing composite

MOS-Specific Rating Scales

10. Overall MOS composite

Combat Performance Prediction Scales

11. Overall Combat Prediction scale composite

Personnel File Form

12. Awards and certificates
13. Disciplinary actions (Articles 15 and Flag actions)
14. Physical readiness
15. Promotion rate

Situational Judgment Test

16. Total composite *or, alternatively,*
17. Discipline soldiers when necessary
18. Focus on the positive
19. Search for underlying causes
20. Work within chain of command
21. Show support/concern for subordinates
22. Take immediate/direct action

Supervisory Simulation Exercises

23. Personal counseling - Communication/Interpersonal skill
24. Personal counseling - Diagnosis/Prescription
25. Disciplinary counseling - Structure
26. Disciplinary counseling - Interpersonal skill
27. Disciplinary counseling - Communication
28. Training - Structure
29. Training - Motivation maintenance

Figure 1.12. Summary List of LVII Basic Criterion Scores

The LVII Data File: Extent of Missing Data

The LVII data were collected from 1,595 soldiers in the nine MOS designated as Batch A MOS in previous data collections. It was not always possible to collect complete information from each soldier for all instruments. For example, for the hands-on measures, some necessary pieces of equipment might have been unavailable for use, making it impossible to score some or all of the steps of a particular task test, or supervisors may have felt that they were not able to use a particular rating scale because of too few opportunities to observe that aspect of performance. For the Personnel File Form, soldiers may have left questions unanswered if they did not know, or chose not to provide, the requested information.

The number of soldiers who are missing all data on a particular instrument can be determined from Table 1.39. For example, only 341 of the 347 MOS 11B soldiers participated in hands-on testing while all 347 soldiers in the 11B sample participated in the job knowledge test administration.

Various methods were used for the criterion instruments to deal with partially missing data. For the Personnel File Form and Simulation Exercises, missing data were simply left as missing. For the other measures, various strategies were used to treat missing data.

The percentages of assigned values for missing data for each performance instrument are shown in Table 1.40. That is, these are the individuals in the sample who had some missing data but not enough to be dropped from the data set for a particular instrument. Instead, their scores were computed using the rules described in Campbell and Zook (1994b). Note that these percentages are generally very low; almost all are less than one percent except for the MOS Ratings Scales.

Development of the LVII Performance Model

The specific objective was to determine which model (i.e., a particular specification of the number of components and their substantive content), from among several proposed alternative models of the latent structure of basic criterion score intercorrelations, best fits the observed data. Analyses were guided by the same general framework that was used in modeling the covariation among performance measures for first-tour performance (J.P. Campbell et al., 1990).

One alternative was the model developed based on data from the Project A Concurrent Validation second-tour (CVII) sample. This model, referred to as the Training and Counseling model, is described in detail in Campbell and Oppler (1990).

Sample and Procedure

The sample used in the LVII modeling analyses included soldiers from eight of the nine Batch A MOS for which a full set of criterion measures had been developed (C.H. Campbell et al., 1990). Because complete data on the entire array of basic criterion scores were required and because soldiers from MOS 31C did not have hands-on performance scores, these soldiers were excluded from all of the present analyses.

Table 1.39

Number of LVII Soldiers With Complete or Partial Data by Criterion Instrument and MOS

MOS	N	Job Knowledge	Hands-On	Army-Wide Rating Scales	MOS Rating Scales	Combat Prediction	PFF	SJT	Simulation Exercises
11B	347	347	341	333	321	313	347	346	341
13B	180	179	174	167	164	160	179	178	174
19K	168	168	160	156	149	152	161	166	156
31C[a]	70	70	--	65	66	64	68	70	--
63B	194	192	187	182	191	176	193	193	188
71L	157	155	156	153	150	150	155	157	156
88M	89	89	88	86	87	85	88	89	88
91A/B	222	220	215	212	208	205	218	220	214
95B	168	168	168	167	162	163	168	168	167
Total	1,595	1,589	1,489	1,521	1,498	1,468	1,577	1,587	1,485

Note. PFF = Personnel File Form; SJT = Situational Judgment Test.

[a] Hands-On and Supervisory Simulation Exercises data were not collected for MOS 31C.

Table 1.40

Percent of LVII Assigned Values by Type of Instrument and MOS

MOS	Job Knowledge	Hands-On	Army-Wide Rating Scales	MOS Rating Scales	Combat Ratings	Personnel File Form	Situational Judgment Test	Supervisory Simulation Exercises
11B	.00	.88	.19	2.88	.00	.00	.14	.00
13B	.00	1.55	.55	2.00	.00	.00	.17	.00
19K	.00	.00	.03	.68	.00	.00	.06	.00
31C	.00	--[a]	.46	1.79	.00	.00	.15	--[a]
63B	.00	1.92	.54	.66	.00	.00	.12	.00
71L	.00	.92	.85	1.75	.00	.00	.11	.00
88M	.00	.91	.44	3.96	.00	.00	.23	.00
91A/B	.00	.92	.78	8.08	.00	.00	.09	.00
95B	.00	.85	.61	6.67	.00	.00	.09	.00
Total Sample	.00	.98	.47	3.33	.00	.00	.12	.00

[a] Hands-On and Supervisory Simulation Exercises data were not collected for MOS 31C.

As a result of these considerations, a total sample of 1,144 soldiers with complete data was available for the initial modeling analyses. The MOS breakdown is shown in Table 1.41.

Table 1.41

Number of LVII Soldiers With Complete Array of Basic Criterion Scores (Excluding Combat Performance Prediction Scales) by MOS

MOS		Number With Complete Data
11B	Infantryman	281[a]
13B	Cannon Crewmember	117
19K	M1 Armor Crewman	105
31C	Single Channel Radio Operator	0
63B	Light-Wheel Vehicle Mechanic	157
71L	Administrative Specialist	129
88M	Motor Transport Operator	69
91A/B	Medical Specialist	156
95B	Military Police	130
Total Sample		1,144

[a] These soldiers do not have general soldiering scores for the hands-on or job knowledge tests.

As a first step, several alternative models of second-tour soldier performance were hypothesized. The fit of these alternative models was then assessed using the LVII data and compared with the fit of the CVII Training and Counseling model. Second, because the Combat Performance Prediction Scales were not included in this initial modeling, key analyses were rerun with these scales included to confirm that the Combat scales fit the models as expected and to determine whether including them would affect the degree of fit. Once a best fitting model was identified, subsequent analyses were conducted to determine whether the model fit equally well across MOS and across demographic subgroups. Finally, based on the results of these analyses, a set of criterion construct scores to be used in the LVII validation analyses was specified.

To generate alternative hypotheses for the latent structure, definitions of the LVII basic criterion scores used in the modeling exercise were circulated to the project staff, and a variety of hypotheses concerning the nature of the underlying structure of second-tour soldier performance were obtained. These hypotheses were consolidated into one principal central alternative model, several variations on this model, and a series of more parsimonious models that involved collapsing two or more of the substantive factors.

The central alternative, the Consideration/Initiating Structure model presented in Table 1.42, differs from the CVII Training and Counseling model primarily in that it includes two leadership factors. Based on staff judgment, the leadership rating scales and each of the SJT and Supervisory Simulation scores were assigned to one of these two factors.

Table 1.42

Consideration/Initiating Structure Model

Latent Variable	Scores Loading on Latent Variables
Core Technical Proficiency (CT)	MOS-Specific Hands-On MOS-Specific Job Knowledge
General Soldiering Proficiency (GP)	General Hands-On General Job Knowledge
Achievement and Effort (AE)	Awards and Certificates Promotion Rate Army-Wide Ratings: Technical Skill/Effort Composite Overall Effectiveness Rating MOS Ratings: Overall Composite Combat Prediction: Overall Composite
Personal Discipline (PD)	Disciplinary Actions (reversed) Army-Wide Ratings: Personal Discipline Composite
Physical Fitness/Military Bearing (PF)	Physical Readiness Score Army-Wide Ratings: Physical Fitness/ Bearing Composite
Leadership: Initiating Structure (IS)	Army-Wide Ratings: Leading/Supervising Composite SE - Disciplinary Structure SE - Counseling Diagnosis/Prescription SE - Training Structure SJT - Disciplining SJT - Immediate/Direct Action SJT - Chain of Command
Leadership: Consideration (LC)	SE - Disciplinary Communication SE - Disciplinary Interpersonal Skill SE - Counseling Communication/Interpersonal Skills SE - Training Motivation Maintenance SJT - Support SJT - Search for Reasons SJT - Focus on the Positive
Written Methods	Technical Knowledge Basic Job Knowledge All Six SJT Scores
Ratings Methods	All Four Army-Wide Ratings Composites Overall Effectiveness Rating MOS Ratings: Overall Composite Combat Prediction: Overall Composite
Disciplinary Simulation Exercise Methods	All Three SE - Disciplinary Counseling Scores
Counseling Simulation Exercise Methods	Both SE - Personal Counseling Scores
Training Simulation Exercise Methods	Both SE - Training Scores

Because the within-MOS sample sizes in the LVII sample were relatively small (ranging from 69 to 281), initial tests of the models were conducted using the entire LVII sample. Criterion scores were first standardized within each MOS, then the intercorrelations among these standardized basic scores were computed across all MOS. The total sample matrix was used as input for the analyses.

The analysis plan was to first compare the fit of the Consideration/ Initiating Structure model with the variations of this model and with the Training and Counseling model to identify the alternatives that best fit the LVII covariance structure. The next set of analyses involved comparing a series of nested models to determine the extent to which the observed correlations could be accounted for by fewer underlying factors. LISREL 7 was used to estimate the parameters and evaluate the fit of each of the alternative models.

Results

The fit of the Training and Counseling model in the LVII sample was remarkably similar to the fit of this same model in the CVII sample, especially considering that the performance data were collected several years apart using somewhat different measures.

Tests of the Consideration/Initiating Structure model and the variations on this model resulted in a very poor fit to the data (e.g., RMSR values greater than .09) and the program encountered a variety of problems in estimating the parameters for these models.

To determine whether there were other reasonable alternative models of second-tour soldier performance, the LVII total sample was randomly divided into two subsamples: 60 percent for model development and 40 percent for cross-validation/confirmation.

The matrix of intercorrelations among the basic criterion scores for the developmental subsample was examined by project staff and several alternative models were suggested. A number of alternatives tried different arrangements of the supervisory simulation, SJT, and rating scale basic scores, while still preserving two leadership factors. None of the these alternatives resulted in a good fit. However, a model that collapsed the Consideration and Initiating Structure factors into a single Leadership factor, included a single Simulation Exercise method factor, and moved the promotion rate variable to the new Leadership factor did result in a considerably better fit to the data in both the developmental and holdout samples.

The "Leadership Factor" model that was developed based on these exploratory analyses is shown in Table 1.43. The fit of the new Leadership Factor model to the LVII data is, for all practical purposes, identical to the fit of the Training and Counseling model to these same data. The 90 percent confidence intervals for the RMSEAs overlap almost completely.

Table 1.43

Leadership Factor Model

Latent Variable	Scores Loading on Latent Variables
Core Technical Proficiency (CT)	MOS-Specific Hands-On MOS-Specific Job Knowledge
General Soldiering Proficiency (GP)	General Hands-On General Job Knowledge
Achievement and Effort (AE)	Awards and Certificates Army-Wide Ratings: Technical Skill/Effort Composite Overall Effectiveness Rating MOS Ratings: Overall Composite Combat Prediction: Overall Composite
Personal Discipline (PD)	Disciplinary Actions (reversed) Army-Wide Ratings: Personal Discipline Composite
Physical Fitness/Military Bearing (PF)	Physical Readiness Score Army-Wide Ratings: Physical Fitness/Bearing Composite
Leadership (LD)	Promotion Rate Army-Wide Ratings: Leading/Supervising Composite SE - Disciplinary Structure SE - Disciplinary Communication SE - Disciplinary Interpersonal Skill SE - Counseling Diagnosis/Prescription SE - Counseling Communication/Interpersonal Skills SE - Training Structure SE - Training Motivation Maintenance SJT - Total Score
Written Method	Job-Specific Knowledge General Job Knowledge SJT - Total Score
Ratings Method	Four Army-Wide Ratings Composites Overall Effectiveness Rating MOS Ratings: Total Composite Combat Prediction: Overall Composite
Simulation Exercise Method	All Seven Simulation Exercise Scores

Because these models have an equally good fit to the data and because the Leadership Factor model does not confound method variance with substantive variance, the Leadership Factor model was chosen as the best representation of the latent structure of second-tour performance.

The Leadership Factor model was tested again with the Combat Performance Prediction Scales included. For one comparison, the Combat Prediction Score was constrained to load only on the Leadership factor and the Rating Method factor. For the second, the Combat Prediction score was constrained to load on the Achievement and Effort and the Rating Method factors only.

The second assignment (i.e., the Combat Prediction Score assigned to the Achievement and Effort factor) produced a much better fit.

__Nested Models__. Next, the Leadership Factor model was used as the starting point to develop a series of more parsimonious nested models, similar to those tested in the LVI sample by Oppler, Childs, and Peterson (1994). The first was identical to the full Leadership Factor model except that the Achievement and Effort factor was collapsed with the Leadership factor.

Similarly, the second nested model was identical to the model just described except that, in addition, the Core Technical and General Soldiering Proficiency factors were replaced with a single "can do" factor. Third, the Personal Discipline factor and the new Achievement/Leadership factor were also collapsed. The fourth model involved adding the variables from the Physical Fitness factor to this Achievement/Leadership/Personal Discipline factor, resulting in a single "will do" factor. The final model collapsed all of the substantive factors into a single overall performance factor.

Because these more parsimonious models are nested within each other, the significance of the loss of fit could be tested by comparing the chi-square values for the various models.

In the first nested model, which involved collapsing the Leadership factor with the Achievement and Effort factor, the resulting decrement in fit was very small. Similarly, collapsing the two "can do" factors resulted in a very small reduction in model fit. Based on these results, a model with only four substantive factors (and three method factors) can account for the data almost as well as the full Leadership Factor model.

Collapsing additional factors beyond this level resulted in larger decrements in model fit.

__Retrospective Re-Analysis of the CVII Data__. One final approach to confirming the Leadership Factor model was to assess the fit of this new model to the CVII data. These results were virtually identical to those obtained in the LVII data.

LVII Criterion Construct Scores

The basic criterion construct scores for use in validation analyses are based on the full Leadership Factor model, with six substantive factors. The nested model with four factors (with a single Achievement/Leadership factor and a single "can do" factor combining Core Technical and General Soldiering Proficiency) fits the data almost as well and has the advantage of greater parsimony. However, it is still plausible that all six performance factors have somewhat different antecedents and could be related to different predictor constructs. Therefore, for the initial validity analyses the model that incorporates the six criterion construct scores was retained.

Results of the nested analyses were used to form more parsimonious sets of criterion construct scores as well. This was done by first standardizing each of the six construct scores described above (based on the full Leadership model). These were then added together in the order shown in Figure 1.13 to form sets of five, four, three, two and finally one criterion composite construct score.

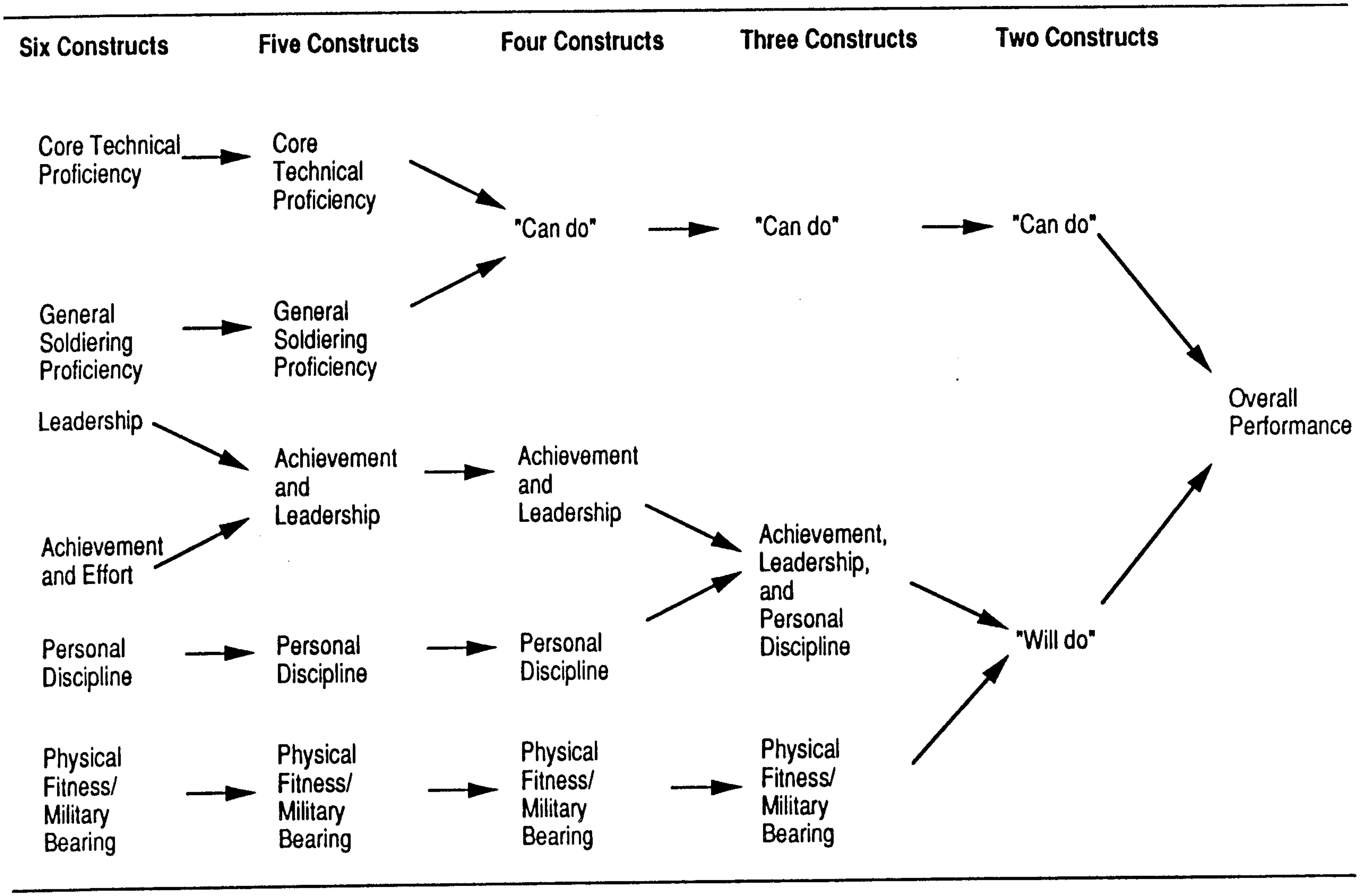


Figure 1.13. Final LVII Criterion and Alternate Criterion Constructs based on more parsimonious models.

Concluding Comments

In general, results of the LVII modeling analyses showed that both the Training and Counseling model and the Leadership Factor model fit the LVII data quite well. Further, retrospective reanalysis of the CVII data showed that these two models had a similarly good fit in the CVII sample.

The new six-factor Leadership Factor model of second-tour performance is also consistent with the CVI/LVI model of first-tour soldier performance. In addition to including performance factors that are parallel to those identified for first-tour soldiers, the LVII second-tour model includes a Leadership factor that contains all measures that were in fact targeted at the leadership/supervision aspects of the job. This is consistent with the results of the second-tour job analyses which indicated that second-tour soldiers perform many of the same tasks as the first-tour soldiers in addition to their supervisory responsibilities. In sum, the Leadership Factor model provides the starting point for the LVII validity analyses and further enhances our understanding of second-tour soldier performance.

ORGANIZATION OF THE CURRENT REPORT

The fourth annual report for the Career Force Project summarizes a number of project activities that have been in the planning stages for some period of time but which were completed during FY 1993. The analysis of attrition using "survival" models and the analysis of results using the Army Job Satisfaction Questionnaire fall in this category. In addition, the Project was called upon to contribute its best efforts to the Joint Service deliberations concerning the forthcoming revision of the ASVAB. Finally, following the specifications for the LVII performance model providing a context for these activities, and described in last year's annual report, the current report describes the basic LVII validation analyses and the consistency of individual soldier performance over time. The chapter organization is as follows:

Chapter 2 describes the Career Force Project's contribution to the deliberations of the Joint Service Manpower Accession Policy Working Group as it developed its proposals for how the ASVAB should be revised, given the available data on the Enhanced Computer-Administered Test Battery (ECAT). The chapter describes an extensive analysis of all possible combinations of subtests relative to variations in time constraints and to changes in the criterion measures.

Chapter 3 reports the basic validation analyses for the available predictors against the LVII performance criterion factors. These analyses parallel those done for CVI, CVII, and LVI. Results are compared across cohorts (1983/84 vs. 1986/87) and across organizational levels.

Chapter 4 presents a brief examination of the consistency of performance from the training context to the first tour and from the first tour to the second tour. Since well-defined performance factors were assessed on the same people at each stage, it is possible to examine the divergent and convergent relationship across factors.

Chapter 5 reports on the Project's use of event history, or survival, analyses to model the prediction of attrition during the first tour of duty. An event history model permits an examination of the time dimension in terms of whether the determinants of attrition are different for different lengths of tenure.

Chapter 6 summarizes an extensive series of analyses using the data gathered with the Army Job Satisfaction Questionnaire (AJSQ). The job satisfaction data were analyzed as a criterion to be predicted, a predictor of attrition and reenlistment/non-reenlistment, and a moderator of the predictor/criterion relationship.

In sum, the fourth annual report will continue the description of the basic validation analyses against job performance that began with the first-tour concurrent validation sample (CVI) and will end with the second-tour longitudinal validation sample. Direct comparisons can now be made across all samples. In addition, an extensive examination of two alternative criteria, attrition and job satisfaction, will be reported. The ECAT analyses will serve as an introduction to the optimal classification battery analyses that still remain to be done, using the entire Experimental Battery and all major performance factors as criterion measures.

Chapter 2

IDENTIFICATION OF OPTIMAL PREDICTOR BATTERIES USING A SUBSET OF THE PROJECT A/CAREER FORCE EXPERIMENTAL PREDICTOR BATTERY

Norman G. Peterson, Christopher E. Sager, Scott H. Oppler, Rodney L. Rosse, and Jennifer L. Crafts

INTRODUCTION

The specific analyses reported in this chapter were completed to assist a subcommittee of the Manpower Accession Policy Working Group in its deliberations about possible revisions to the ASVAB. The committee had previously identified a battery of experimental tests, designated as the Joint Services Enhanced Computer Administered Test battery (ECAT), that would constitute the array of new predictors to be considered as potential additions to ASVAB. Data were collected across Services and analyzed by several sets of analysts (Kieckhaefer et al., 1992; Wolfe et al. (in preparation); Abrahams et al., 1993; Peterson, Oppler, Sager, & Rosse, 1993). The analyses reported here are limited to the experimental predictors in the ECAT battery that were also in the Project A/Career Force Experimental Battery. The Joint Services are also looking at Navy tests.

The practical problem is to decide which subset of the ten ASVAB subtests plus seven tests from the Project A/Career Force predictor battery should be included in an improved operational test battery. This task is complicated by the fact that a number of indices can be used to evaluate the performance of a test battery. The indices considered in this chapter include (a) the validity of the battery, (b) the battery's capacity to maximize classification efficiency, (c) the differences that could result from use of the test battery across multiple subgroups (i.e., Blacks, Hispanics, and Females), (d) the performance of the battery relative to these first three parameters across multiple MOS, and (e) the amount of time required to administer the battery. The following sections discuss some recent studies in which each of these indices of test battery performance has been considered.

Indices for Evaluating the Value of a Test Battery

Validity

A recent assessment of the validity of the ASVAB is reported by Oppler, McCloy, and Peterson in Chapter 3 of the Career Force 1991 Annual Report (Campbell & Zook, 1994a). Across 22 U.S. Army jobs, the mean multiple correlation for predicting end-of-training school knowledge scores with scores on the ASVAB subtests was .77. Note that (a) the multiple correlations were adjusted for shrinkage and corrected for range restriction before they were averaged, (b) the criteria were scored on written tests that were developed separately for each job, and (c) while the ASVAB consists of 10 subtests, analyses performed for the Army routinely combine the scores on the Word Knowledge and Paragraph Comprehension subtests to form a single Verbal score, resulting in a total of nine subtest scores.

Oppler et al. also reported the incremental validities of the Career Force Experimental Battery of predictors. As summarized in the preceding chapter, the Experimental Battery was developed to measure (a) cognitive

abilities not already thoroughly assessed by the ASVAB, (b) psychomotor abilities, (c) temperament constructs hypothesized to be related to job performance, and (d) interest constructs hypothesized to be related to job performance. The results indicate that while the Experimental Battery did result in incremental validity when predicting ratings of training performance, it did not show incremental validity for the prediction of end-of-training school knowledge.

In this chapter, absolute validity is used as the index of validity. Absolute validity is defined here as the multiple correlation associated with the regression of a criterion on a set of predictors.

Classification Efficiency: Differential Validity

One way of assigning applicants to jobs would be to (a) use a single set of subtest weights to generate a predicted performance score for each applicant, (b) select all applicants who score above some cutoff on this predicted score, and (c) randomly distribute the selected applicants across jobs. This situation will be referred to here as selection without classification. An alternative method of assigning applicants to jobs would be to (a) generate a unique predicted performance score for each applicant in each job and (b) assign applicants to the various jobs (or to the Not Selected group) in a manner that maximizes the mean predicted performance of the applicants. This situation will be referred to here as classification. Conceptually, classification efficiency can be viewed as the increase in the average level of job performance that would result from replacing the first method of assignment (selection without classification) with the second method of assignment (classification).

A necessary condition for classification efficiency is differential validity. In the situation where a unique least-squares regression equation is calculated for each job, differential validity is the degree to which the subtest weights in these equations vary across jobs. Differential validity results in intra-individual variation in predicted performance scores across jobs. With this intra-individual variation, a different score would serve to predict an individual's performance in each job. Therefore, there would be a potential advantage to optimally matching applicants to jobs relative to random assignment after the select/not select decision is made. (This explanation assumes that there are not meaningful mean differences in performance across jobs. See Johnson & Zeidner (1990) for an explanation of the potential classification efficiency associated with mean differences in performance across jobs.)

Using a sample of 78,041 enlistees in 82 U.S. Air Force jobs, Ree and Earles (1991) assessed differential validity in the prediction of training-school grades with the 10 subtests of the ASVAB. Their results indicated that, after controlling for mean differences in grades across jobs, using unique subtest weights for each job resulted in a multiple correlation of .63; using only one set of subtest weights for all jobs resulted in a multiple correlation of .61. Ree and Earles inferred from this result that ". . . a common equation for all jobs was almost as predictive as an equation for each job" (p. 321). A single equation means no intra-individual variation in predicted performance across jobs and, therefore, no potential classification efficiency.

Our index of differential validity will be referred to as discriminant validity and defined as the difference between mean absolute validity and mean generalizability validity. Mean absolute validity is the average multiple correlation resulting from the regression of the criterion on a set of predictors in each of multiple jobs. That is, the predicted performance scores for each job are computed using regression weights that are specific to that job. The predicted performance scores from the equation using the least-squares weights computed within each job can also be correlated with performance scores in each of the other jobs. The mean of these correlations is the mean generalizability validity. If there are (m) jobs, this mean is based on m(m-1) validity estimates.

The difference between mean absolute validity and mean generalizability validity is an assessment of the extent to which there is intra-individual variation in predicted performance across jobs. Thus, our index of differential validity is:

$$\text{Discriminant Validity (DV)} = \text{mean absolute validity} - \text{mean generalizability validity}$$

We will use this index of potential classification efficiency in this chapter because of its close intuitive relationship with differential validity. It has also been used in an investigation of validity in a prior Army Research Institute project (Wise, Peterson, Hoffman, Campbell, & Arabian, 1991).

Classification Efficiency: Brogden Index of Classification Efficiency

Johnson and Zeidner (1990) argue that differential validity is less than optimal as an indicator of potential classification efficiency. They advocate using a more direct estimate of the benefit derived from optimally assigning applicants to jobs. One such index -- Brogden's mean predicted performance (MPP) (Brogden, 1959) -- is obtained by calculating

$$R(1-r)^{1/2}\ f(A)$$

where R is the mean of the absolute validities across jobs, r is the mean of the correlations among the predicted scores for each job, and f(A) is a multiplier that increases with the selection ratio and with the number of jobs for which classification occurs (see Peterson et al., 1993), for a more detailed discussion of the equation used to compute mean predicted performance and the components that make up the equation).

The formula estimates the effect of actually completing the classification process (e.g., assigning each individual to one job) and then calculating the mean of the predicted performance scores. The advantages of the equation are that it (a) simultaneously considers indicators of absolute validity and differential validity, and (b) takes into account two important contextual selection variables -- the number of jobs and the selection ratio. Several authors (Brogden, 1959; Johnson & Zeidner, 1990; Statman, 1993) have shown significant increases in potential classification efficiency associated with relatively small increases in absolute validity and differential validity.

However, several disadvantages are associated with using Brogden's MPP. There are underlying assumptions that may not hold in all contexts. Two that are not realistic assumptions for the Project A/Career Force data are that

there is (a) no variation in R across jobs (the least-squares regression equations yield the <u>same</u> absolute validities across all jobs), and (b) no variation in r across jobs (the same degree of similarity exists among all pairs of equations).

For the purposes of this report, we will explore the use of an index of predicted performance and evaluate its performance against our indices of absolute and differential validity. However, we use a modified version of Brogden's mean predicted performance and refer to it as the Brogden index of classification efficiency:

$$R(1-r)^{1/2}$$

where, as before, R is the mean of the absolute validities across jobs and r is the mean of the correlations among the predicted scores for each job. The multiplier f(A) has been eliminated from this modified index because the number of jobs and the selection ratio do not vary; thus, the Brogden index of classification efficiency values will not order a set of batteries differently than will the Brogden MPP values.

There are an important similarity and an important difference between differential validity and the Brogden index. The similarity is that R and mean absolute validity are the same value. The difference is that mean generalizability validity is the mean of the correlations between the observed performance scores in each job and the predicted performance scores based on the equations developed in all of the other jobs, whereas r is the mean of the correlations among predicted performance scores only (i.e., predicted scores based on the equations developed in each job are correlated with predicted scores based on the equations developed in each of the other jobs).

Johnson and Zeidner (1990) suggest that computer simulations of assignments, based on algorithms designed to maximize mean predicted performance, currently provide the most accurate estimate of the potential benefit to be gained from classification. Predicted performance scores based on empirical or synthetic data are calculated for each individual in each job; then an algorithm assigns individuals to jobs in some optimal fashion, and the mean predicted performance is calculated based on each individual's predicted performance score in the job to which he or she was assigned.

The authors point out that (a) this method of estimating potential classification efficiency does not depend on the restrictive assumptions underlying Brogden's MPP, and (b) various constraints imposed on the assignment algorithm (e.g., a different quota for each job) do not reduce the accuracy of the estimate of potential classification efficiency. While it is true that Ree and Earles (1991) demonstrated that the ASVAB subtests produce a relatively small amount of differential validity, Statman (1993) used this simulation method to demonstrate that using the ASVAB subtests to assign individuals to jobs would result in an increase in mean predicted performance relative to the Army's current use of aptitude area composites. The primary disadvantage of this method is that performing a computer simulation for each possible battery of subtests is extremely resource intensive and virtually impossible to complete for the kinds of analyses undertaken here.

In summary, we will use two indices of potential classification efficiency in this chapter, the discriminant validity and the Brogden index of classification efficiency, and make appropriate comparisons of the results.

Subgroup Differences

The Uniform Guidelines on Employee Selection Procedures (Equal Opportunity Employment Commission, 1978) that were adopted by the federal government urge that (a) organizations show that predictors currently being used to select and classify job applicants treat subgroups fairly, and (b) the domain of individual differences used to predict job performance be expanded. The latter admonition is based on the hypothesis that subgroup differences might be reduced by the discovery and use of predictors that demonstrate smaller subgroup differences. In response to these recommendations, the military has expended considerable effort to establish the fairness of the ASVAB subtests (e.g., Wise et al., 1992) and to search for and develop new predictors of job performance (e.g., Peterson, Hough, et al., 1990).

In this chapter, we examine subgroup differences between White and Black (W-B), White abd Hispanics (W-H), and Males and Females (M-F). Subgroup differences will be operationally defined as the mean of the predicted scores for individuals in the referent subgroup (e.g., Whites) minus the mean of the predicted scores for individuals in the non-referent subgroup (e.g., Blacks) divided by the standard deviation of the referent subgroup's predicted scores.

Issues in Optimizing the Usefulness of Test Batteries

In addition to the above, a number of other issues need to be addressed when forming batteries from a selection of subtests. That is, what method should be used to form the alternative battery, and what trade-offs will be faced in situations where batteries perform differentially on the targeted indices?

Results for a Single Index of Battery Quality

Research on absolute validity suggests that it is reasonable to expect that the ECAT battery subtests might provide the ASVAB with incremental validity (McHenry et al., 1990). Research on classification suggests that subtests which measure more specific cognitive abilities are likely to increase the differential validity of the ASVAB, and that relatively small increases in absolute validity and differential validity can have substantial impact on potential classification efficiency (Johnson & Zeidner, 1990). Research on subgroup differences suggests that adding measures of more specific cognitive abilities to the ASVAB may decrease the level of differences associated with the battery (Jensen, 1985; Scholarios, 1990). This research does not, however, emphasize the possibility that adding a particular subtest to the ASVAB might reduce one type of subgroup differences but increase another type.

Evaluation of Predictor Batteries Using Multiple Indices

Broadening the domain of cognitive abilities currently assessed by the ASVAB could simultaneously increase the battery's level of absolute validity and/or differential validity and decrease the level of subgroup differences (e.g., Scholarios, 1990; Johnson & Zeidner, 1990). The rationale for this

prediction is based on three observations: (a) General cognitive ability accounts for a substantial portion of the variance in the ASVAB's current subtests (Ree & Earles, 1991), (b) measures of specific cognitive abilities tend to show smaller subgroup differences than measures of general cognitive ability (Jensen, 1985), and (c) measures of specific cognitive abilities would increase the absolute and differential validity of the current ASVAB (McHenry et al., 1990; Johnson & Zeidner, 1990).

However, as noted earlier, a potential problem not emphasized in prior research is that measures of specific abilities that reduce differences between one set of subgroups might increase them for another set. For example, adding a measure of a particular psychomotor ability might reduce differences between White and Blacks, but increase them between Males and Females.

Absolute Validity Versus Classification Efficiency

How does one choose an "optimal" battery if the combination of subtests that maximizes absolute validity is not the same as the combination of subtests that maximizes potential classification efficiency? Brogden (1959) indicates that substantial gains in classification efficiency can be achieved with relatively small increases in absolute validity and differential validity; however, there are trade-offs between these two parameters (Statman, 1993). If all the assumptions are met, Brogden's MPP can determine which combination of subtests would result in the greatest mean predicted performance. However, because Brogden's assumptions are rarely met (and are not met by the data analyzed in this report), his equation for estimating MPP is of limited usefulness for choosing among alternative batteries that maximize absolute validity versus maximizing differential validity.

Johnson and Zeidner's method of computer simulations is capable of addressing this issue (Scholarios, 1990). Scholarios based a computer simulation on the Project A data set. The results of the simulation indicated that when the criterion was core technical proficiency and the predictors were the ASVAB subtests and the Project A experimental test battery, test batteries that were chosen based on maximizing differential validity resulted in higher mean predicted performance than test batteries that were chosen based on maximizing absolute validity. It is important to note that these simulations involved relatively few potential combinations of subtests, as compared to the number that will be considered in this chapter.

An Alternative Method for Identifying "Optimal" Batteries

A variation on a stepwise regression method was used by Abrahams, Pass, Kusulas, Cole, and Kieckhaefer (1993) to identify optimum combinations of 10 ASVAB and 9 ECAT subtests to maximize absolute validities of batteries ranging in length from 1 to 19 subtests.

For each analysis, the steps were as follows: (a) The single test that had the highest weighted average validity across schools (jobs) was selected, (b) multiple correlations were computed for all possible combinations of the first test with each of the remaining tests, and (c) the average multiple correlation was then computed for each test pair and the combination with the highest weighted average was selected. This process was repeated until all remaining tests were included. The idea is that at each step, the subtests

currently in the prediction equation represent the test battery of that length that maximizes average absolute validity.

Three limitations are associated with this method. First, because this method does not allow subtests to be dropped at later stages of accretion (i.e., the procedure uses forward stepwise regression only), the particular combination of subtests at the n^{th} step may not be the battery of n tests that maximizes mean absolute validity. The relative contribution of a particular subtest may diminish as other subtests are added to the battery.

Second, this method seeks only to identify the optimal battery relative to the mean absolute validity across jobs; it does not account for potential classification efficiency or differences among subgroups. One implication of this difficulty is that for a test battery of any length, the combination of subtests identified by this method may produce a relatively low level of differential validity and relatively high subgroup differences compared to another combination of subtests with a similar level of mean absolute validity, but a higher level of differential validity and lower subgroup differences.

A final limitation of this method is that while the number of subtests in a test battery is related to the cost of administering the battery, the actual test battery administration time might provide a more accurate assessment of cost.

<u>Summary of Issues</u>

The goal of the analyses reported here is to identify an "optimal" combination of tests selected from the ASVAB subtests and Project A/Career Force ECAT predictors for predicting three criteria (technical/school knowledge, on-the-job core technical proficiency, and on-the-job hands-on test performance) at various time lengths for test battery administration. In the context of this study, "optimal" is defined as the maximization of three distinct types of battery performance indices.

However, there are likely to be trade-offs involved when building batteries. There is no established way to choose between two batteries when one battery has greater absolute or differential validity and another battery has smaller subgroup differences, or between two batteries when one decreases one type of subgroup difference (e.g., W-B) and the other decreases another type (e.g., M-F). Yet, if the relative standings of alternative batteries on the various test battery performance indices are known, then it may be possible to make explicit judgments about which battery to use.

A NEW APPROACH

It is not possible to choose a combination of subtests that simultaneously optimizes all the relevant indications of test battery quality. It is possible, however, to calculate all the indices of test battery performance for every possible combination of subtests that fall within a given test administration time interval. For example, this method can be used to compute the absolute validity, differential validity, Brogden's index, W-B, W-H, and M-F differences of every combination of subtests that require from 134 to 164 minutes of administration time. These combinations of subtests can

then be rank-ordered according to each index of test battery performance. For instance, the top 20 test batteries ranked on the basis of maximum absolute validity can be compared to the top 20 test batteries ranked on the basis of minimum M-F differences.

An advantage of this method is that it provides explicit information necessary to evaluate trade-offs: If subtests are included to optimize test battery performance on one index, how will the battery perform on the other indices?

This approach was one of those followed in the Joint Services Enhanced Computer Administered Test validation study designed to evaluate different combinations of 19 ASVAB and ECAT subtests in predicting end-of-training performance (Peterson et al., 1993). Data were collected from 9,037 Air Force, Army, and Navy enlistees representing 17 military jobs. The analysis procedures used three time intervals, included two ASVAB subtests (Arithmetic Reasoning and Word Knowledge) in every potential test battery in each of the three time intervals, and evaluated absolute validity, differential validity, classification efficiency, and three types of subgroup differences (W-B, W-H, and M-F) for each battery in each time interval.

Results indicated that no single test battery (within each time interval) simultaneously optimized all the test battery indices examined. The researchers identified trade-offs associated with maximizing absolute validity or classification efficiency vs. minimizing all three types of subgroup differences, and with minimizing M-F differences vs. minimizing either W-B or W-H differences.

The same approach was used to obtain the results reported in this chapter. However, the analyses here were done for each of three different Project A/Career Force crtierion measures and the sample was limited to Army personnel only.

Objectives

Two primary objectives were addressed in the analyses reported in this chapter. The first was to examine the battery indices for each potential test battery at three time intervals that are much shorter, somewhat shorter, and approximately the same length as the current ASVAB. These lengths were chosen pragmatically; it is unlikely that additional time can be made available for military testing, but shorter time periods could be viewed as desirable. The second objective was to explore the trade-offs associated with choosing test batteries that individually optimize each of the six parameters of test battery performance.

We used an approach similar to that described for the ECAT study above (Peterson et al., 1993). The same analysis procedure was followed; we used three distinct battery time intervals and included the Arithmetic Reasoning and Word Knowledge ASVAB subtests in every potential test battery for each of the three time intervals.

Sample

The data were from nine of the Longitudinal Validation (LV) first-tour Batch A MOS (19E was excluded); they are listed in Table 2.1. Recall that these data were collected using a longitudinal validation design with approximately two years between the time when predictor tests were administered to soldiers in the sample and the time when criterion data were collected on their job performance. The end-of-course training performance technical knowledge test was administered 3-6 months after the individual started basic training.

Table 2.1

MOS Sample Sizes for Calculating Absolute, Generalizability, and Discriminant Validities of Test Batteries

	Criterion		
MOS	Technical/School Knowledge	On-the-Job Core Technical Proficiency	On-the-Job Hands-On Performance
11B	2,210	292	292
13B	3,356	696	696
19K	1,555	543	543
31C	564	206	206
63B	936	510	510
71L	1,257	305	305
88M	1,084	273	273
91A	2,956	622	622
95B	3,401	312	312

Measures

Criteria

Three LV first-tour criteria were used in these analyses. The first was Technical/School Knowledge, an end-of-training criterion. The other two were on-the-job criteria: Core Technical Proficiency and Hands-On Test Performance. Sample sizes for each of the three criterion measures are listed in Table 2.1. A total of 17,319 soldiers were tested for the Technical/School Knowledge criterion, and 3,759 were tested for the criteria of Core Technical Proficiency and Hands-On Test Performance.

Predictors

Our Career Force analyses were designed to identify potential combinations of 15 candidate predictors to be administered along with the Arithmetic Reasoning and Word Knowledge subtests of the ASVAB. These two subtests were designated as the base condition; as such, they were included in each of the potential test batteries discussed below. The 15 predictors considered included all the remaining ASVAB subtests and the Experimental

Battery subtests under serious consideration for supplementing the ASVAB. Specifically, the tests were:

- Eight ASVAB subtests other than Arithmetic Reasoning (ASAR) and Word Knowledge (ASWK):
  - -- Auto and Shop Information (ASAS)
  - -- Coding Speed (ASCS)
  - -- Electronics Information (ASEI)
  - -- General Science (ASGS)
  - -- Mechanical Comprehension (ASMC)
  - -- Mathematical Knowledge (ASMK)
  - -- Numerical Operations (ASNO)
  - -- Paragraph Comprehension (ASPV)

- Three Project A/Career Force paper-and-pencil spatial tests:
  - -- Assembling Objects (SPAO)
  - -- Orientation (SPOR)
  - -- Reasoning (SPRS)

- Three Project A/Career Force computerized tests:
  - -- Target Identification (Decision Time) (CMTI)
  - -- One-Hand Tracking (CM1T)
  - -- Two-Hand Tracking (CM2T)

- One Project A/Career Force computerized test composite:
  - -- Short-Term Memory (CMST)

These subtests are described in greater detail in Peterson, Russell, et al. (1990).

Test Battery Time Intervals

Three different test battery time intervals were related for this study. These intervals were defined by total minutes allowed for "start-up" time, instruction time, and test-taking time. The three time intervals were: 74-104 minutes, 134-164 minutes, and 194-224 minutes. Start-up time was specified to last 20 minutes, regardless of the number or types of tests to be included in the battery, and instruction time was set at 3 minutes per test. The amount of time required to take each test, not including instructions, is listed in Table 2.2.

Analyses

The following analyses were conducted for each of the three criterion measures and three time intervals specified above.

Absolute Validity

The multiple correlation that was associated with all possible combinations of predictors for which the total test administration time was within the specified time interval, and that included the Arithmetic Reasoning and Word Knowledge subtests, was computed within each MOS. For example, there are 10,152 unique combinations of predictors for which the total test administration time is between 134 and 164 minutes (and which include the Arithmetic Reasoning and Word Knowledge subtests). Therefore, 10,152 multiple

Table 2.2

Amount of Time Required for Each Test

Test	Title	Source	Length (minutes)[a]
ASAR[b]	Arithmetic Reasoning	ASVAB	36
ASWK[b]	Word Knowledge	ASVAB	11
ASAS	Auto and Shop Information	ASVAB	11
ASCS	Coding Speed	ASVAB	7
ASEI	Electronic Information	ASVAB	9
ASGS	General Science	ASVAB	11
ASMC	Mechanical Comprehension	ASVAB	19
ASMK	Mathematical Knowledge	ASVAB	24
ASNO	Numerical Operations	ASVAB	3
ASPC	Paragraph Comprehension	ASVAB	13
SPAO	Assembling Objects	Spatial	18
SPOR	Orientation	Spatial	10
SPRS	Reasoning	Spatial	12
CMST	Short-Term Memory	Computer	5
CMTI	Target Identification	Computer	3
CM1T	One-Hand Tracking	Computer	4
CM2T	Two-Hand Tracking	Computer	4

[a] Does not include 3 minutes on each test for instructions.
[b] Included in all test batteries.

correlations were computed within each MOS (a different multiple correlation for each combination of predictors meeting the above requirements).

These multiple correlations were corrected for multivariate range restriction using the Lord and Novick (1968) formula and the 10 x 10 ASVAB covariance matrix from the 1980 Youth Population (Department of Defense, 1982). They were also adjusted for shrinkage using Rozeboom's (1978) Formula 8. The unweighted average multiple correlations across the nine MOS for each combination are referred to as the mean absolute validities.

Potential Classification Efficiency: Discriminant Validity

The least-squares equation computed within each MOS (for each possible combination) was also used to predict performance in each of the other MOS. That is, for each unique test combination, each of the nine MOS equations was applied to the other eight MOS. This resulted in the computation, for each unique test battery, of 72 correlations between predictor equations developed in one MOS and performance in a different MOS. Note that these correlations were not adjusted for shrinkage (because there is no capitalization on chance).

The unweighted average of these 72 correlations for each combination is referred to as the mean generalizability validity, and the difference between the mean absolute validity and the mean generalizability validity is referred

to as the mean discriminant validity. This latter index provides an indication of the extent to which the regression weights associated with a particular combination of predictors vary across MOS.

Potential Classification Efficiency: Brogden Index

The Brogden index for the nine MOS was calculated for each combination of subtests that included Arithmetic Reasoning and Word Knowledge, as follows:

$$\text{Mean absolute validity} \times [1 - \text{mean correlation among the predicted scores}]^{1/2}$$

Subgroup Differences

Finally, three estimates of subgroup differences (i.e., White vs. Black; White vs. Hispanic; Male vs. Female) were computed for each possible combination of predictors. This was done by applying the least-squares regression equations computed within each MOS (discussed above) to all soldiers in the LV predictor sample. (The LV predictor sample included 25,538 whites, 9,058 blacks, 1,252 Hispanics, 32,393 males, and 4,722 females.) Each equation generated a composite score for each soldier.

Next, mean composite scores were calculated by subgroup for each equation. Estimates of subgroup differences were computed for each equation by subtracting the non-referent subgroup (e.g., Blacks) mean composite score from the referent subgroup (e.g., Whites) mean composite score and dividing this value by the standard deviation of the referent subgroup's composite scores. These estimates were then averaged across equations, resulting in three indices of subgroup differences (i.e., White vs. Black; White vs. Hispanic; Male vs. Female). An advantage of this approach is that the estimates of subgroup differences were relatively stable because they were calculated on a relatively large sample.

RESULTS

The results are summarized in Table 2.3. The separate sections of the table present the general summary statistics and information about the performance of the prediction equations across all possible combinations of predictors (that each include Arithmetic Reasoning and Word Knowledge) for each time interval and for each of the criterion measures.

The first three rows report: (a) the number of possible combinations of tests for each time limit, (b) the mean (and standard deviation) of the testing times associated with the combinations for each time limit, and (c) the mean (and standard deviation) of the number of tests associated with those combinations. For example, the number of combinations of tests that fit into the 134-164 minute interval is 10,152. The average amount of time associated with these combinations is 151.0 (SD = 8.7) minutes, and the average number of tests is 8.2 (SD = 1.0).

Average Battery Performance Indices for Each Criterion

The remainder of Table 2.3 contains the average score on each of the six predictor battery performance indicators for each of the three criteria and each time interval. The six indicators are (a) the mean absolute validity

Table 2.3

Summary of Results for Predicting Each Criterion at Each Time Interval

General Summary Statistics	74-104 Minute Batteries		134-164 Minute Batteries		194-224 Minute Batteries	
Total Number of Combinations	210		10,152		6,509	
Mean Testing Time in Minutes (SD)	97.5	(5.8)	151.0	(8.7)	205.7	(8.4)
Mean Number of Tests (SD)	4.6	(.7)	8.2	(1.0)	11.6	(1.0)
Technical/School Knowledge						
Mean Validity Coefficient (SD)	.739	(.009)	.761	(.007)	.771	(.005)
Mean Discriminant Validity Index (SD)	.013	(.008)	.025	(.007)	.031	(.004)
Mean Brogden Index (SD)	.102	(.026)	.145	(.018)	.164	(.010)
Mean W-B Difference (SD)	1.426	(.054)	1.457	(.045)	1.465	(.028)
Mean W-H Difference (SD)	1.053	(.042)	1.080	(.034)	1.088	(.023)
Mean M-F Difference (SD)	.262	(.089)	.315	(.069)	.321	(.044)
On-the-Job Core Technical Proficiency						
Mean Validity Coefficient (SD)	.593	(.010)	.615	(.009)	.623	(.006)
Mean Discriminant Validity Index (SD)	.012	(.007)	.022	(.007)	.027	(.005)
Mean Brogden Index (SD)	.123	(.018)	.163	(.015)	.186	(.010)
Mean W-B Difference (SD)	1.339	(.063)	1.378	(.047)	1.392	(.029)
Mean W-H Difference (SD)	.993	(.051)	1.019	(.043)	1.023	(.030)
Mean M-F Difference (SD)	.342	(.114)	.407	(.092)	.411	(.057)
On-the-Job Hands-On Test Performance						
Mean Validity Coefficient (SD)	.455	(.017)	.490	(.012)	.500	(.007)
Mean Discriminant Validity Index (SD)	.017	(.007)	.027	(.007)	.030	(.006)
Mean Brogden Index (SD)	.132	(.014)	.167	(.012)	.187	(.010)
Mean W-B Difference (SD)	1.182	(.103)	1.197	(.065)	1.193	(.035)
Mean W-H Difference (SD)	.888	(.083)	.899	(.055)	.895	(.034)
Mean M-F Difference (SD)	.526	(.175)	.636	(.125)	.656	(.074)

W-B = White-Black; W-H = White-Hispanic; M-F = Male-Female.

coefficient, (b) the mean discriminant validity index, (c) the mean Brogden's index of classification efficiency, (d) the mean White vs. Black difference, (e) the mean White vs. Hispanic difference, and (f) the mean Male vs. Female difference.

For example, the average validity for predicting Technical/School Knowledge was .739 (SD = .009) across the 210 test combinations fitting into the 74-104 minute time interval. The average discriminant validity was .013 (SD = .008); the average Brogden's index was .102 (SD = .026); the average White vs. Black difference was 1.426 (SD = .054); the average White vs. Hispanic difference was 1.053 (SD = .061); and finally the average Male vs. Female difference was .262 (SD = .089).

Correlations Among Performance Indicators for Each Criterion for Potential Batteries at Each Time Interval

Tables 2.4 through 2.12 contain the correlations among the mean predictor battery performance indicator scores for each criterion and each time interval. For example, Table 2.8 shows that when predicting On-the-job Core Technical Proficiency at the 134-164 minute time interval (across the 10,152 combinations of predictors), the average correlation (across the nine MOS) between discriminant validity and White vs. Black differences is .36. This means that, for this criterion/time interval, combinations of predictors with relatively low discriminant validity tended to have relatively low White vs. Black differences, and combinations of predictors with relatively high discriminant validity tended to have relatively high White vs. Black differences.

Table 2.4

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 210 Combinations for Technical/School Knowledge: 74-104 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.88	1.00				
Brogden Index	0.87	0.99	1.00			
W-B Difference	0.48	0.51	0.45	1.00		
W-H Difference	0.41	0.35	0.31	0.86	1.00	
M-F Difference	0.78	0.87	0.88	0.35	0.24	1.00

W-B = White-Black; W-H = White-Hispanic; M-F = Male-Female

Table 2.5

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 10,152 Combinations for Technical/School Knowledge: 134-164 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.88	1.00				
Brogden Index	0.88	1.00	1.00			
W-B Difference	0.66	0.55	0.53	1.00		
W-H Difference	0.51	0.42	0.42	0.69	1.00	
M-F Difference	0.75	0.87	0.87	0.48	0.45	1.00

Table 2.6

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 6,509 Combinations for Technical/School Knowledge: 194-224 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.92	1.00				
Brogden Index	0.92	1.00	1.00			
W-B Difference	0.65	0.57	.56	1.00		
W-H Difference	0.44	0.39	.39	0.60	1.00	
M-F Difference	0.73	0.86	.84	0.50	0.48	1.00

Table 2.7

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 210 Combinations for On-the-Job Core Technical Proficiency: 74-104 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.56	1.00				
Brogden Index	0.43	0.94	1.00			
W-B Difference	0.61	0.35	0.17	1.00		
W-H Difference	0.48	0.09	-0.04	0.85	1.00	
M-F Difference	0.57	0.61	0.56	0.44	0.27	1.00

Table 2.8

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 10,152 Combinations for On-the-Job Core Technical Proficiency: 134-164 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.46	1.00				
Brogden Index	0.41	0.93	1.00			
W-B Difference	0.83	0.36	0.27	1.00		
W-H Difference	0.38	0.08	0.04	0.61	1.00	
M-F Difference	0.44	0.75	0.67	0.45	0.37	1.00

Table 2.9

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 6,509 Combinations for On-the-Job Core Technical Proficiency: 194-224 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.48	1.00				
Brogden Index	0.35	0.91	1.00			
W-B Difference	0.75	0.24	0.12	1.00		
W-H Difference	0.18	-0.06	-0.10	0.49	1.00	
M-F Difference	0.43	0.73	0.60	0.36	0.36	1.00

Table 2.10

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 210 Combinations for On-the-Job Hands-On Test Performance: 74-104 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.67	1.00				
Brogden Index	0.56	0.92	1.00			
W-B Difference	0.26	0.33	0.11	1.00		
W-H Difference	0.21	0.26	0.10	0.88	1.00	
M-F Difference	0.67	0.71	0.61	0.19	0.04	1.00

Table 2.11

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 10,152 Combinations for On-the-Job Hands-On Test Performance: 134-164 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.57	1.00				
Brogden Index	0.47	0.91	1.00			
W-B Difference	0.69	0.32	0.19	1.00		
W-H Difference	0.37	0.25	0.19	0.65	1.00	
M-F Difference	0.51	0.63	0.51	0.32	0.24	1.00

Table 2.12

Correlations Among Mean Predictor Battery Performance Indicator Scores Across 6,509 Combinations for On-the-Job Hands-On Test Performance: 194-224 Minute Interval

Indicators	Val.	Dis. V	BI	W-B	W-H	M-F
Validity	1.00					
Discriminant Validity	0.58	1.00				
Brogden Index	0.38	0.92	1.00			
W-B Difference	0.68	0.22	0.10	1.00		
W-H Difference	0.27	0.18	0.14	0.50	1.00	
M-F Difference	0.42	0.49	0.37	0.22	0.21	1.00

Battery Rankings for Performance Indices

Appendix A presents results listing the top 20 potential test batteries identified for each of the six test battery performance indices. These results are reported separately for each of the nine sets of analyses (Tables A1 through A9). That is, results are given for each time interval within each of the three criterion measures. Each table presents a particular criterion/time interval combination; each table has six lists of equations rank ordered by a performance index.

The first list (i.e., page) in each table is the 20 combinations of predictors with the greatest mean absolute validity across the nine MOS. These 20 combinations are itemized under the heading NV Predictor Variables. Also shown for each combination are the mean discriminant validity, the Brogden index, the mean White vs. Black difference estimate, the mean White vs. Hispanic difference estimate, the Male vs. Female difference estimate, and the required test administration time associated with that combination of tests. (The columns for these indices are labeled Validty Est (Mean and S.D.), Discr Vldty, Brog. Index, Whte-Blck (Mean and Rnge), Whte-Hisp (Mean and Rnge), Male-Fem (Mean and Rnge), and Test Time, respectively.)

The second through sixth lists (i.e., pages) in each Appendix A table follow the same format to present, in turn, the top 20 combinations of predictors with respect to discriminant validity, the Brogden's index, the minimization of White vs. Black differences, the minimization of White vs. Hispanic differences, and the minimization of Male vs. Female differences, respectively.

Note that the last row in each listing reports the results associated with a test battery solely consisting of the ASVAB Arithmetic Reasoning and Word Knowledge subtests (labeled Base).

DISCUSSION

This section discusses the results and addresses the issues associated with identifying optimal predictor batteries for each criterion measure. The following questions direct the discussion for each criterion measure:

1. Across the three battery time intervals, what is the average battery value on each of the six indices?

The average value of each index for all the potential test batteries at each time interval is summarized in each of the subsections titled "Mean Values on Battery Performance Indices."

2. How do the mean values for the six indices change as battery length increases?

We will examine potential gains or trends in the indices as the battery length is extended in the subsections titled "Changes in Index Performance With Increasing Battery Length."

3. What trade-offs are associated with optimizing performance on one index vs. the other indices? In other words, what are the correlations among the six indices?

The magnitudes of the relationships among the test battery indicators are discussed in order to illustrate the trade-offs among indices that may be faced in situations where batteries must be formed from a set of available subtests. As an example, one battery could maximize the absolute validity criterion but not also simultaneously minimize one or more of the specific indices of subgroup differences. This subsection is titled "Relationships Among Indices" in the discussion for each criterion.

4. Are certain subtests (or type of subtests, e.g., computer-administered) favored in the top 20 rankings for a specific index or across a variety of indices? That is, will forming a battery by optimizing alternative indices favor the same or different subtests?

The relative frequencies with which the various subtests are included in the listings of the top 20 batteries are mentioned in the subsections, "Frequencies of Subtest Occurrence in Rankings."

5. How do the two indices of classification efficiency relate -- do they order tests similarly or differently?

We will examine the degree of agreement between these two indices in terms of the frequency with which each subtest is included in the index rankings. These comparisons are discussed in the subsections "Discriminant Validity vs. the Brogden Index."

Predicting Technical/School Knowledge

Mean Values on Battery Performance Indices

The average level of absolute validity (mean validity coefficient) predicting Technical/School Knowledge across the three time intervals is substantial (about .74 - .77; see Table 2.3). The variability of the validities is relatively small (i.e., the standard deviations are < .01). The average discriminant validities range from .013 to .031. The average Brogden Index is about .137.

The average estimates of White vs. Black differences are slightly less than one and a half standard deviations. The average estimates of White vs. Hispanic are slightly more than a standard deviation. Finally, the estimates of Male vs. Female differences are between one-quarter and one-third of a standard deviation.

Changes in Index Values With Increasing Battery Length

Relatively small increases in validity, either absolute or discriminant, occur as the length of the battery is extended (see Table 2.3). For example, the mean validity coefficient for the prediction of Technical/School Knowledge increases by about .03 as the battery length is increased from the 74-104 minutes range to the 194-224 minutes range. Similarly, the increase in discriminant validity between the shortest and longest time interval is less than .02. The Brogden index increases by about .06 as the battery time is lengthened. There seems to be a very slight increase in subgroup differences as the time interval lengthens; however, the increases from the shortest to the longest time interval are less than one-tenth of a standard deviation.

Relationships Among Indices

Tables 2.4 through 2.6 show the relationships among the equation performance indices for Technical/School Knowledge. There is a substantial positive relationship between absolute validity and discriminant validity (.88, .88, and .92 for the short, moderate, and long intervals). This means that combinations of tests with relatively high mean absolute validities are

likely to also have relatively high mean discriminant validities. The relationships between absolute validity and the Brogden index are of the same strength as the absolute validity - discriminant validity relationships.

The correlations between absolute validity and W-B differences and between discriminant validity and W-B differences are relatively substantial for the three time intervals. This means that combinations of subtests with relatively high mean absolute validities and relatively high mean discriminant validities are very likely to have relatively high levels of W-B differences. The correlations of absolute validity and discriminant validity with W-H differences are lower than with W-B differences for all three time intervals.

The relationships of absolute validity and discriminant validity with M-F differences are the strongest of all (in the .78 - .87 range); these correlations drop off slightly as the time interval increases. An important point is that the M-F indices of M-F differences have stronger positive relationships with discriminant validity than with absolute validity. We will address this issue in more detail later. The three indicators of subgroup differences are positively related. However, another way in which M-F differences are unique is that they are more weakly related to both W-B and W-H subgroup differences than the W-B and W-H indices are to each other.

Frequencies of Subtest Occurrence in Rankings

Information about predictor combinations on Technical/School Knowledge is shown in Appendix Tables A1, A2, and A3 for the short, moderate, and long time intervals, respectively. For all three intervals, the top 20 equations in terms of absolute and discriminant validity all include the ASVAB subtest Auto and Shop Information. The computer tests are also common in these equations.

The spatial tests occur much more frequently in the rankings for minimizing subgroup differences than for maximizing either absolute validity or discriminant validity. Also for the top 20 batteries that minimize each of the three types of subgroup differences, the ASVAB subtest Coding Speed shows up consistently and Auto and Shop Information does not appear. At all three intervals, the ASVAB subtest Numerical Operations consistently appears in the batteries that minimize W-B and W-H differences; at the two longer intervals, Mathematical Knowledge also appears consistently.

Discriminant Validity vs. the Brogden Index

Tables 2.4 through 2.6 show that there is complete agreement between the two indices of classification efficiency. The correlations are .99 for the shortest interval and 1.00 for the other two intervals. Although it is no surprise that they are strongly correlated, the magnitude of the relationships is higher than expected.

Predicting On-the-Job Core Technical Proficiency

Mean Values on Battery Performance Indices

The average level of absolute validity (mean validity coefficient) predicting On-the-Job Core Technical Proficiency across the three time intervals is substantial (about .59 - .62, see Table 2.3). The variability of

the validities is relatively small (i.e., the standard deviations are < .01). The average discriminant validities range from .012 to .027 across the three time intervals. The average Brogden indices range from .123 to .186.

The average estimates of White vs. Black differences are slightly above one and a third standard deviations. The average estimates of White vs. Hispanic differences are approximately one standard deviation. Finally, the estimates of Male vs. Female differences are between three- and four-tenths of a standard deviation.

Changes in Index Values With Increasing Battery Length

There are relatively small increases in validity, either absolute or discriminant, as the length of the battery is extended (see Table 2.3). The mean validity coefficient for the prediction of On-the-job Core Technical Proficiency increases by .03 as the battery length is increased from the 74-104 minute range to the 194-224 minute range, and the increase in discriminant validity between the shortest and longest time interval is less than .02. Again, as for the Technical/School Knowledge criterion, there is an increase of about .06 in the Brogden index.

There seems to be a general, but very slight, increase in subgroup differences between the first two time intervals and then a smaller increase in subgroup differences between the medium interval and the long time interval. The increase from the short to the medium time interval is in no case greater than one- tenth of a standard deviation.

Relationships Among Indices

The relationships among the equation performance indices for On-the-Job Core Technical Proficiency are shown in Tables 2.7 through 2.9. For all three time intervals, there is a moderate positive relationship between absolute validity and discriminant validity. This means that combinations of tests with relatively high mean absolute validities are likely to have relatively high mean discriminant validities. Correlations of absolute validity with the Brogden index are slightly lower than with discriminant validity for all three time intervals. Correlations between discriminant validity and the Brogden index range from .94 for the short time interval to .91 for the longest time interval.

The correlations of absolute validity with W-B differences are relatively substantial. This means that combinations of tests with relatively high mean absolute validities are very likely to have relatively high levels of W-B differences. Correlations between discriminant validity and W-B differences are slightly more than half the magnitude of the corresponding correlations between absolute validity and W-B differences.

The relationship between absolute validity and W-H differences is relatively weaker, and drops off from .48 for the short time interval to only .18 for the longest time interval. The correlations between discriminant validity and W-H differences are very low, indicating almost no relationship between these indices.

The relationship between absolute validity and M-F differences is moderate (ranging from . 57 for the shortest to .43 for the longest interval).

However, the relationship between discriminant validity and M-F differences is more substantial (approximately three-tenths of a standard deviation at the medium and longest time intervals). The three indices of subgroup differences are positively related: W-B and W-H are more strongly related to each other than either is to M-F differences, with the lowest correlations between W-H and M-F over the three time intervals.

Frequencies of Subtest Occurrence in Ranking

Information about predictor combinations of On-the-Job Core Technical Proficiency is shown in Appendix Tables A4, A5, and A6 for the 74-104, 134-164, and 194-224 minute time intervals, respectively. The top 20 battery rankings for absolute validity and discriminant validity all include (with two exceptions in the shortest interval) the ASVAB subtest Auto and Shop Information. At the two longer time intervals the ASVAB subtest Coding Speed is also in every equation. The Short-Term Memory computer subtest is also common in these equations. In the two longer time intervals, the spatial tests appear more frequently in the validity battery rankings, whereas the computer tests occur more frequently in the discriminant validity battery rankings.

The Numerical Operations test does not show up in the rankings for validity, discriminant validity, or minimizing M-F differences, but does show up consistently in the rankings for minimizing both W-B and W-H differences. Auto and Shop Information does not appear in the rankings for minimizing each of the three types of adverse impact, but the ASVAB subtests Coding Speed and Mathematical Knowledge do show up relatively frequently. The difference between the batteries that minimize W-B and W-H differences and the batteries that minimize M-F differences is that the former include a fair number of computer tests but the latter include very few computer tests.

Discriminant Validity vs. the Brogden Index

As shown in Tables 2.7 through 2.9, the Brogden index and discriminant validity are again very strongly related, but not to the same degree as for the Technical/School Knowledge criterion. For the shortest interval, the correlation is .94; this decreases to .93 and .91 as the time interval lengthens.

Predicting On-the-Job Hands-On Test Performance

Mean Values on Battery Performance Indices

The average level of absolute validity (mean validity coefficient) predicting On-the-Job Hands-On Test Performance across the three time intervals is relatively moderate (about .46 - .50; see Table 2.3). The variability of the validities is relatively small (i.e., the standard deviations are < .02). The average discriminant validities range from .017 to .030. The average Brogden indices range from .132 to .187.

The average estimates of White vs. Black differences are approximately one and a fifth standard deviations. The average estimates of White vs. Hispanic differences are around nine-tenths of a standard deviation. Finally, the estimates of Male vs. Female differences are between five-tenths and two-thirds of a standard deviation.

Changes in Index Values With Increasing Battery Length

As shown in Table 2.3, there are relatively small increases in validity, either absolute or discriminant, as the length of the battery is extended. The mean validity coefficient for the prediction of On-the-Job Hands-On Test Performance increases by about .04 as the time interval is increased from the shortest to the longest battery length. Once again, the increase in discriminant validity is less than .02, and the increase in the Brogden index between the shortest and longest time interval is about .06. There is very little change in subgroup differences over the three time intervals for the W-B and W-H indices, but a more marked increase of .13 standard deviation over these intervals for the M-F index.

Relationships Among Indices

The relationships among the equation performance indices for On-the-Job Hands-On Test Performance are shown in Tables 2.10 through 2.12. There is a relatively moderate positive relationship between absolute validity and discriminant validity. This means that combinations of tests with relatively high mean absolute validities are likely to have relatively high mean discriminant validities. This holds for all three time intervals. In comparison, absolute validity correlates less strongly with the Brogden index. The correlations between discriminant validity and the Brogden index are .92, .91, and .92 for the short, moderate, and long intervals, respectively.

The correlations between absolute validity and W-B differences vary greatly -- from .26 to .69 and .68 as the battery length increases. For the longer two intervals only, this means that combinations of tests with relatively high mean absolute validities are very likely to have relatively high levels of W-B differences. This means that the chances of discovering combinations of tests with relatively low levels of W-B differeances and relatively high absolute validities are better for the shorter time interval. The correlations between absolute validity and W-H differences and discriminant validity and W-H differences are relatively low.

The relationships between absolute validity and M-F differences and discriminant validity and M-F differences for the shortest interval are moderate (.67 and .71, respectively) but decrease as the time interval increases. Again, the index of M-F differences is the only indicator of subgroup differences that has a stronger positive relationship with discriminant validity than with absolute validity. Again, the three indices of subgroup differences are positively related; however, W-B and W-H differences are more related to each other than either is to M-F differences.

Frequencies of Subtest Occurrence in Rankings

Appendix Tables A7, A8, and A9 contain the battery rankings for predicting On-the-Job Hands-On Test Performance for the short, moderate, and long time intervals, respectively. As for the two criteria discussed previously, the absolute and discriminant validity rankings all include the ASVAB subtest Auto and Shop Information. The computer tests are also common in these equations (especially Target Identification, One-Hand Tracking, and Two-Hand Tracking). Also, the computer tests show up more frequently than the spatial tests in the top 20 absolute and discriminant validity equations.

Auto and Shop Information does not appear in the top 20 subtest combinations that minimize each of the three types of subgroup differences, but the ASVAB subtests Coding Speed and Mathematical Knowledge show up relatively often. Again, the test batteries that minimize W-B and W-H differences include the Numerical Operations subtest; this subtest is missing from the rankings for minimizing M-F differences at the two shorter intervals but does appear in about half of the top 20 batteries for the longest time interval. The batteries that minimize both W-B and W-H differences also include a moderate number of computer tests; the batteries that minimize M-F differences favor the spatial tests over the computer tests.

Discriminant Validity vs. the Brogden Index

For On-the-Job Hands-On Test Performance, correlations between the Brogden index and discriminant validity are .92, .91, and .92 for the 74-104, 134-164, and 194-224 minute intervals, respectively. These correlations are of the same magnitude as those for the other on-the-job criterion of Core Technical Proficiency.

Classification Efficiency: Discriminant Validity vs. the Brogden Index

As described in the sections for each criterion measure above, there is a considerable overlap between discriminant validity and the Brogden index of classification efficiency. It is not a surprise that the correlations for all criteria and time interval combinations are above .90 because both indices are measures of classification efficiency. These results indicate that the two indices order the potential test batteries very similarly, especially for the Technical/School Knowledge criterion. A less predictable result is that for the Technical/School Knowledge criterion, the correlations between absolute validity and the Brogden index are essentially equivalent to those between absolute validity and discriminant validity, but for the two On-the-Job criteria the correlations of absolute validity with the Brogden index are lower than correlations of absolute validity with discriminant validity.

Tables A1 through A9 in Appendix A show that there is substantial agreement between the two indices of potential classification efficiency in terms of the frequency with which each subtest is included in each index's list of top 20 test batteries. There is, however, a consistent difference between the rankings for these two indices across all tables. On the average, more variables are listed for each of the equations for the Brogden index than for each of the equations for discriminant validity. Despite this difference, the two indices function very similarly. Consequently for the remainder of this chapter, we will focus on the index of discriminant validity.

"Optimal" Test Batteries

Any number of potential rules could be used to try to create an "optimal" test battery that somehow considers all the discussed indices of battery performance. As a demonstration, two such rules were implemented using the "top twenty" potential test batteries in Appendix A.

Two Potential "Rules" for Identifying Battery Subtests

The first "rule" implemented was "Maximize Validity/Minimize Average Subgroup Differences." According to this rule, the optimal combination of tests for each criterion and time interval was that battery of tests in the top 20 mean absolute validities list (see Appendix A) that also had the lowest average across the three types of subgroup differences. (This rule considered only the top 20 absolute validity equations.) The optimal equations according to this rule for each criterion and time interval are presented in Table 2.13.

According to these optimal equations, there is some improvement in absolute validity, discriminant validity, and the Brogden index between the short time interval and the medium time interval; however, the improvement in most cases is less dramatic between the medium time interval and the long time interval.

The mean W-B difference ranges from 1.2 to 1.4 standard deviations for all the criterion/time interval combinations. The W-H difference is approximately one standard deviation for Technical/School Knowledge and On-the-Job Core Technical Proficiency, but appears to be somewhat smaller for On-the-Job Hands-On Test Performance. The M-F difference is less than one-half of a standard deviation for Technical/School Knowledge and On-the-Job Core Technical Proficiency, but is substantially greater for On-the-Job Hands-On Test Performance.

Beyond the automatically included Arithmetic Reasoning and Word Knowledge, Auto and Shop Information is in almost every optimal equation. The computer-administered memory test shows up in all the longer batteries. The computer-administered tracking tests appear in the longest battery for School Knowledge and the intermediate-length battery for Hands-On, not at all for Core Technical Performance.

The second "rule" implemented was "Maximize Discriminant Validity/ Minimize Average Subgroup Difference." According to this rule, the optimal subtest combination for each criterion and time interval was that battery of tests in the top 20 mean discriminant validities list (see Appendix A) that also had the lowest average across the three types of subgroup differences. (Again, this rule considered only the top 20 batteries ranked according to the discriminant validity index.) The optimal equations according to this rule for each criterion and time interval are presented in Table 2.14.

The results in Table 2.14 are in most respects very similar to the results in Table 2.13. A major difference is that the computer tests occur more often in these "optimal" equations, especially for the longer intervals and for those equations predicting On-the-Job Hands-On Test Performance. This result suggests that the computer tests contribute more to maximizing discriminant validity. The problem is, however, that they also appear to contribute to M-F differences. There is a general tendency for the estimates of M-F differeances in Table 2.14 to be of greater magnitude than the corresponding estimates in Table 2.13. This trade-off between discriminant validity and M-F differences is also demonstrated by the substantial correlations between discriminant validity and M-F differences in Tables 2.4 through 2.12 (especially Tables 2.10, 2.11, and 2.12).

Table 2.13

"Optimal" Test Batteries for Each Criterion and Time Interval According to the "Maximize Validity/Minimize Average Subgroup Differences Rule

Criterion (Time Interval)	V	DV	BI	W-B	W-H	M-F	T (min.)	Test Battery[a]
TSK (74-104 mins.)	.759	.029	.153	1.448	1.037	0.356	103	ASAR,ASWK,ASAS,ASCS,ASNO
TSK (134-164 mins.)	.774	.033	.167	1.453	1.057	0.346	161	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMST,CMTI
TSK (194-224 mins.)	.779	.036	.177	1.447	1.068	0.348	218	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO, SPOR,CMST,CMTI, CM1T, CM2T
CTP (74-104 mins.)	.610	.006	.101	1.410	1.106	0.352	100	ASAR,ASWK,ASEI,SPRS
CTP (134-164 mins.)	.633	.029	.171	1.413	0.993	0.417	153	ASAR,ASWK,ASAS,ASCS,ASMK,SPAO,CMST
CTP (194-224 mins.)	.633	.034	.191	1.397	0.994	0.439	203	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC, SPAO,CMST
HOT (74-104 mins.)	.486	.027	.149	1.164	0.857	0.725	101	ASAR,ASWK,ASAS,ASNO,CMST
HOT (134-164 mins.)	.513	.034	.177	1.210	0.873	0.682	164	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CM1T
HOT (194-224 mins.)	.514	.034	.189	1.237	0.933	0.597	200	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS, CMST,CMTI

TSK = Technical/School Knowledge; CTP = On-the-Job Core Technical Proficiency; HOT = On-the-Job Hands-On Test Performance; V = Validity; DV = Discriminant Validity; BI = Brogden Index of Classification Efficiency; W-B = White - Black Difference; W-H = White - Hispanic Difference; M-F = Male - Female Difference; T = Testing Time for Battery.

[a] See Table 2.2 for the titles and sources of the individual tests.

Table 2.14

"Optimal" Test Batteries for Each Criterion and Time Interval According to the "Maximize Discriminant Validity/Minimize Average Subgroup Differences" Rule

Criterion (Time Interval)	V	DV	BI	W-B	W-H	M-F	T (min.)	Test Battery[a]
TSK (74-104 mins.)	.759	.029	.153	1.448	1.037	0.356	103	ASAR,ASWK,ASAS,ASCS,ASNO
TSK (134-164 mins.)	.770	.036	.173	1.433	1.071	0.393	157	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
TSK (194-224 mins.)	.778	.036	.177	1.449	1.069	0.348	220	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASPC, SPAO,CMST,CMTI,CM1T
CTP (74-104 mins.)	.615	.029	.158	1.386	0.990	0.478	103	ASAR,ASWK,ASAS,ASCS,ASNO
CTP (134-164 mins.)	.621	.039	.200	1.360	0.980	0.519	163	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,ASPC,CMST,CM1T,CM2T
CTP (194-224 mins.)	.622	.041	.211	1.357	0.990	0.502	203	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
HOT (74-104 mins.)	.461	.029	.161	1.028	0.785	0.596	100	ASAR,ASWK,ASNO,CMST,CMTI,CM2T
HOT (134-164 mins.)	.494	.044	.201	1.139	0.856	0.794	161	ASAR,ASWK,ASAS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
HOT (194-224 mins.)	.509	.043	.210	1.183	0.894	0.668	214	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T

TSK = Technical/School Knowledge; CTP = On-the-Job Core Technical Proficiency; HOT = On-the-Job Hands-On Test Performance; V = Validity; DV = Discriminant Validity; BI = Brogden Index of Classification Efficiency; W-B = White - Black Difference; W-H = White - Hispanic Difference; M-F = Male - Female Difference; T = Testing Time for Battery.

[a] See Table 2.2 for the titles and sources of the individual tests.

Contributions of the ASVAB Subtests

Validity and subgroup differences were evaluated at the test battery level because the influence that a particular subtest has on the test battery performance indices depends on the other subtests in that particular battery. However, it is still useful to observe the frequency with which subtests occur in the top 20 lists according to each test battery index. These observations afford a general sense of the tendency for a particular subtest to optimize a test battery's contribution with respect to particular indices.

This section summarizes the subtests selected for the 134-164 minute time interval. This is the time interval that allows the greatest differential inclusion of subtests across batteries. No reference is made to either the Arithmetic Reasoning or Word Knowledge subtests of the ASVAB because these were designated as the base condition (automatically included in every potential test battery).

Auto and Shop Information contributes to absolute validity and discriminant validity, but not to minimizing all three types of subgroup differences. Coding Speed contributes to minimizing all three types of subgroup differences. Electronic Information tends to make less of a contribution to absolute validity and more of a contribution to discriminant validity. General Science tends to make more of a contribution to absolute validity and much less of a contribution to discriminant validity, and Mechanical Comprehension tends not to contribute to either absolute or discriminant validity.

Mathematical Knowledge tends to contribute to absolute validity to a lesser extent and to discriminant validity to a greater extent. Mathematical Knowledge also has a tendency to contribute to minimizing W-B differences but has a greater tendency to contribute to minimizing W-H and M-F differences. Numerical Operations has (a) a low tendency to contribute to absolute validity and discriminant validity, (b) a high tendency to contribute to minimizing W-B and W-H differences, and (c) a low tendency to contribute to minimizing M-F differences. Paragraph Comprehension has a high tendency to contribute to minimizing W-B and M-F differences, and a low tendency to contribute to minimizing W-H differences.

Adding Subtests to the ASVAB

Another question is "What is gained by adding subtests to the existing set of ASVAB subtests?" Assembling Objects tends to contribute to both absolute validity and the minimization of M-F differences. Orientation tends not to contribute to any indices, and Reasoning tends to contribute to minimizing M-F differences.

Target Identification has a strong tendency to contribute to discriminant validity and a moderate tendency to contribute to minimizing all three types of differences. One-Hand Tracking has a low to moderate tendency to contribute to absolute validity and a high tendency to contribute to discriminant validity. One-Hand Tracking tends to contribute more to absolute validity and discriminant validity than Two-Hand Tracking does. One-Hand and Two-Hand Tracking have a moderate tendency to contribute to minimizing W-B and W-H differences but not to minimizing M-F differences.

Substituting Subtests for the ASVAB Subtests

The issue of whether Project A/Career Force subtests might be substituted for one or more of the ASVAB subtests can be approached from several different perspectives, such as time allotted for test administration, and feasibility of administering different types of tests (e.g., computer capabilities), as well as maximizing validity indices or minimizing subgroup differences.

The 194-224 minute interval most closely approximates the length of the current ASVAB. For this interval, for the Technical/School Knowledge criterion, Appendix Table A3 shows that all ASVAB subtests except Mechanical Comprehension, Coding Speed, and Numerical Operations always appear in the rankings for absolute validity and discriminant validity. This means that a number of spatial and computer tests tended to contribute more to maximizing these two indices than did these three ASVAB tests. Therefore, only seven or eight of the ASVAB subtests (including Arithmetic Reasoning and Word Knowledge), rather than all ten, were included in the batteries that included 12 to 14 subtests. On the other hand, for the rankings according to subgroup differences, the subtests that were in almost all cases excluded from the batteries were Auto and Shop Information, Electronic Information, and Mechanical Comprehension. These subtests were left out of the equations because spatial and computer subtests contributed more to minimizing subgroup differences.

For the 194-224 minute interval for the On-the-Job Core Technical Proficiency criterion, the subtests that appeared less frequently for the absolute and discriminant validity indices were Numerical Operations and General Science (see Table A6). For the three indices of subgroup differences, the subtests Auto and Shop Information, Electronic Information, and Mechanical Comprehension failed to appear in the top 20 batteries.

The results obtained for the 134-164 minute interval are applicable to situations where less testing time would be available. For this interval for the Technical/School Knowledge criterion (see Table A2), the Coding Speed, Mechanical Comprehension, and Numerical Operations subtests were included least often of all ASVAB subtests in the battery rankings for the absolute and discriminant validity indices. Again as for the longest interval, the Auto and Shop Information, Electronic Information, and Mechanical Comprehension subtests as well as the General Science subtest were excluded from the top 20 rankings for the three indices of subgroup differences.

For the On-the-Job Core Technical Proficiency criterion at the 134-164 minute level (see Table A5), the General Science, Mechanical Knowledge, and Numerical Operations subtests were the most often excluded in the rankings for absolute and discriminant validity. For all three indices of subgroup differences, four subtests were excluded from the rankings: Auto and Shop Information, Electronic Information, Mechanical Comprehension, and General Science.

At the 74-104 minute interval, so few subtests are included in the batteries that it is not very informative to consider which ASVAB subtests do not occur in the rankings. Of course, the patterns mentioned above for the longer intervals also hold for the shortest interval, and are magnified due to the time constraint.

Assuming a situation where ASVAB subtests not included in the rankings were to be dropped from an operational battery, then testing time could be filled with subtests from the Project A/Career Force battery that did contribute to the indices. These subtests were described, corresponding to the indices they contributed to, in the section "Adding Subtests to the ASVAB," above.

Comparisons Across Criteria

The validity coefficients presented to this point were not corrected for criterion unreliability. This correction is necessary to make comparisons across criteria. Although the content of the criterion measures differs, these measures were developed and administered in very similar ways across MOS. For purposes of approximation, we corrected the mean validity coefficients in Table 2.3 using the criterion reliability estimates listed in Table 2.15. Table 2.15 provides the mean validity coefficients corrected for unreliability. These results are consistent with past research -- when the predictors are ability measures, on-the-job performance criteria (i.e., Core Technical Proficiency and Hands-On Test Performance) tend to be somewhat less predictable than end-of-training performance criteria (i.e., Technical/School Knowledge).

Table 2.15

Mean Validity Coefficients Corrected for Unreliability for Each Criterion and Time Interval

Criterion	R_{yy} [a]	74-104 Mins.	134-164 Mins.	194-224 Mins.
Technical School/Knowledge	.90	.779	.802	.812
On-the-Job Core Technical Proficiency	.85	.638	.662	.671
On-the-Job Hands-On Test Performance	.50	.644	.693	.707

[a] Reliability estimates for each criterion are based on data from Campbell (1987); estimates are slightly higher than observed values so that conservative corrections can be made.

Concluding Remarks

The method of analysis we used (i.e., calculating the test battery performance indices for every possible combination of subtests at predetermined time intervals) does not identify the combination of subtests that will simultaneously optimize all the test battery performance parameters considered

in this report.[1] The results of these analyses do, however, convincingly make the point that such a test battery does not exist. Examination of the test batteries ordered by each of the test battery performance indices makes some trade-offs apparent.

The more dramatic trade-offs are (a) the maximization of absolute validity versus the minimization of all three types of subgroup differences, (b) the maximization of classification efficiency versus the minimization of all three types of subgroup differences, (c) the minimization of W-B differences versus the minimization of M-F differences, and (d) the minimization of W-H differences versus the minimization of M-F differences.

Predictably, subtests in the ASVAB battery show a substantial tendency to contribute to absolute validity and a moderate tendency to contribute to potential classification efficiency. Compared to the ASVAB subtests, the Project A/Career Force subtests show a relatively more substantial tendency to contribute to classification efficiency and a relatively moderate tendency to contribute to absolute validity. The Project A/Career Force subtests do not seem to show a dramatically greater tendency, compared to the ASVAB subtests, to contribute to minimizing the three types of subgroup differences.

It is important to note that there are a number of reasons why a particular subtest can be included in a top 20 test batteries list according to a particular test battery performance index. First, all the test batteries examined here automatically included Arithmetic Reasoning and Word Knowledge; therefore, the contributions of the other subtests to optimizing each index were relative to these two subtests. The reason is that the contribution of every subtest to the optimization of each index is relative to the other subtests in that particular potential test battery.

The reasons for inclusion also depend on the index. If the index is mean absolute validity, a subtest that accounts for a unique portion of the variance in the criterion across all jobs will appear in many of the top 20 test batteries. If the index is discriminant validity, a subtest that accounts for a unique portion of the variance in the criterion in one or more of the jobs but not in other jobs will appear in many of the top 20 test batteries.

Subtests can also appear in the top 20 test batteries ranked according to the minimization of a particular type of subgroup difference (e.g., W-B differences) for multiple reasons. One way for a subtest to contribute to a relatively small subgroup difference on the predicted performance score is for it to predict a portion of the variance in the criterion on which there is a relatively small subgroup difference. However, a subtest's contribution to a relatively small subgroup difference on the predicted performance score may be independent of the subtest's validity or the size of the subgroup difference on that subtest. This is because the subtest's contribution to the predicted performance score is a function of the regression weight that is assigned to it by the least-squares regression equation.

[1] This method was employed with a set of tests that had been carefully chosen and for which criterion-related validity had already been demonstrated; we do not recommend this technique for purely exploratory validation analyses.

The point is that a subtest that receives a small regression weight, for whatever reason, will have very little influence on the predicted performance score and will not increase the subgroup difference on the score substantially. Therefore, a subtest might be included in a potential test battery that has a relatively small subgroup differences because that subtest has a small influence on the predicted performance score but helps the battery reach the predetermined time interval.

As an example, assume that: (a) two subtests require the same amount of time to administer, (b) the first subtest, upon which Males score higher than Females, would increase the absolute validity of the test battery substantially, and (c) the second subtest is unreliable and not related to the criterion, but Males and Females do not score differently on it. All else being equal, the minimization of M-F differences will favor a battery that includes the second subtest over a battery that includes the first subtest.

An important point is that we imposed the condition that each test battery fit into one of three predetermined administration time intervals. Therefore, a combination of subtests may have ranked in the top 20 test batteries for a particular index not because it performed well on that index, but because it performed well relative to the other potential test batteries that also fit into that time interval.

A Suggested Approach -- Conducting Utility Analyses for a Narrowed Set of Batteries

In terms of evaluating the value of a particular test battery utility, mean absolute validity and discriminant validity are good indices for ordering equations within a time interval; however, mean absolute validity and discriminant validity may not provide the best possible information about the relative utility of shorter versus longer batteries. To make this type of inference, the approach described here could be combined with the MPP approach proposed by Johnson & Zeidner (1990). The steps would be to (a) identify some number of top equations for each time interval, for each of the indices we have used in this chapter, and (b) calculate MPP using computer simulations to generate an estimate of the potential job performance gain associated with specific test batteries.

Implications of Weighing Alternative Goals in Complying With Professional and Legal Guidelines

Value judgments will be faced in any situation where a battery of subtests is formed for operational use. We have described some of the indices to consider in these situations. If judgments regarding the relative priority of optimizing each of the test battery parameters are not made explicitly, they will have been made implicitly by default.

The results reported here have important implications for organizations attempting to comply with professional and legal guidelines regarding test development and use. The Uniform Guidelines stipulate that employers search for alternatives among available predictors that are job-related; these alternatives must at the same time have less adverse impact against protected subgroups. Our analyses suggest that the subtests included in batteries selected to minimize differences among subgroups are not equivalent to those included in batteries selected to maximize absolute (or discriminant)

validity. They further indicate that the subtests included in batteries selected to minimize differences between one set of subgroups often differ substantially from the subtests included in batteries selected to minimize differences between a different set of subgroups.

Chapter 3
BASIC VALIDATION RESULTS FOR THE LVII SAMPLE

Scott H. Oppler, Norman G. Peterson, and Andrew M. Rose

This chapter summarizes the results of the validity evaluation of the ASVAB and the Project A Longitudinal Validation (LV) Experimental Battery for predicting second-tour performance in the Army. The results are based on the second-tour performance data collected from the Project A/Career Force longitudinal sample. They extend the validation results reported previously (Oppler, Peterson, & Russell, 1994) for first-tour (LVI) performance.

OBJECTIVES

The objectives of the analyses described in this chapter are as follows:

(1) Compute the basic validities for ASVAB and Experimental Battery predictors against the second-tour performance factors and selected individual performance measures.

(2) Compare the validities of four alternative sets of ASVAB scores (nine ASVAB subtests vs. four ASVAB factors vs. AFQT vs. MOS-appropriate Aptitude Area composites).

(3) Compare the validities of three alternative sets of ABLE scores.

(4) Assess the incremental validities for the Experimental Battery predictors over the four ASVAB factor composites.

(5) Compare the incremental validities of three alternative sets of ABLE scores.

(6) Compare the validities and incremental validities of the Experimental Battery predictors reported for LVI with the validities and incremental validities reported for LVII.

PROCEDURES

Sample

The results reported here are based on a different strategy for editing the LVII sample than was previously used for the LVI analyses. To review briefly, the LVI analyses used two sampling plans, a "listwise deletion" plan and a "setwise deletion" plan. The first of these strategies mirrored the procedure used to evaluate the Project A CVI predictors against first-tour performance. To be included in those analyses, soldiers in the CVI sample were required to have complete data for all of the Project A CV predictor composites, as well as for the ASVAB and all of the CVI first-tour performance factors. Corresponding to this CV strategy, a validation sample composed solely of soldiers having complete data for all the LVI Experimental Battery predictors, the ASVAB, and the LV first-tour performance factors was created for the LVI data set. This sample was referred to as the "listwise deletion" sample.

In the alternative LVI sample editing strategy, a separate validation sample was identified for each set of predictors in the Experimental Battery. More specifically, to be included in the validation sample for a given predictor set, soldiers were required to have complete data for each of the first-tour performance factors, the ASVAB, and the predictor composites in that predictor set only. For example, a soldier who had data for the complete set of ABLE composites (as well as complete ASVAB and criterion data), but was missing data from the AVOICE composites, would have been included in the "setwise deletion" sample for estimating the validity of the (ABLE) test, but not in the sample for the AVOICE test.

The principal reason for creating these setwise deletion samples was to maximize the sample sizes used to estimate the validity of the Experimental Battery predictors. The setwise sample sizes were, in fact, considerably larger than those associated with the listwise strategy.

The number of soldiers with complete predictor and criterion data (the listwise deletion sample) in each MOS is reported in Table 3.1 for both the LVI and LVII samples. Note that in the LVII sample only two MOS (63B and 91A) have at least 100 soldiers with complete predictor and criterion data. Also, there are no soldiers with complete data in the LVII sample for MOS 31C; no MOS-specific criteria or supervisory simulation data were administered to soldiers in that MOS, so those soldiers were not included in the LVII validation analyses.

Table 3.1

Soldiers in LVI and LVII Data Sets With Complete Predictor and Criterion Data by MOS

MOS		LVI	LVII
11B	Infantryman	235	83
13B	Cannon Crewmember	553	84
19K	M1 Armor Crewman	446	82
31C[a,b]	Single Channel Radio Operator	172	--
63B	Light-Wheel Vehicle Mechanic	406	105
71L	Administrative Specialist	252	77
88M[b]	Motor Transport Operator	221	37
91A/B	Medical Specialist	535	118
95B	Military Police	270	93
Total		3,163	679

[a] MOS-specific and supervisory simulation criterion data were not collected for MOS 31C in LVII.
[b] MOS 31C and 88M were not included in LVII validity analyses.

The number of soldiers in each MOS in the LVII sample who were able to meet the setwise deletion requirements used in the LVI analyses is shown in Table 3.2. Even this relaxed sample selection strategy results in three MOS (19K, 71L, and 88M) with sample sizes of consistently fewer than 100 soldiers.

Table 3.2

Soldiers in LVII Sample Meeting Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	277	269	97	251	248	257
13B	116	112	102	107	97	105
19K	103	89	90	84	88	88
63B	152	132	124	123	117	126
71L	126	101	92	93	96	91
88M[a]	66	46	43	43	43	44
91A	155	143	133	139	128	137
95B	127	119	110	114	105	115
Total	1,122	1,011	791	954	922	963

[a] MOS 88M was not included in LVII validity analyses.

As a result of this finding, a third sample selection strategy, which we have termed "predictor/criterion setwise deletion," was considered to enlarge the samples for the present analyses. This strategy extends the reasoning underlying the setwise strategy employed in the LVI analyses. Specifically, to be included in the validation sample for a given criterion/predictor set pair, soldiers were required to have complete data for the ASVAB, the predictor composites in the predictor set being examined, and only the specific criterion score being predicted. Thus, in addition to not requiring that a soldier have complete predictor data to be included in a particular validity analysis, the present strategy also did not require that the soldier have complete data for all of the criterion scores.

As expected, this strategy increases the sample sizes available for estimating the validity of the Experimental Battery predictors against second-tour performance. Table 3.3 reports the number of soldiers in each MOS meeting the predictor/criterion setwise deletion sample criteria for the Core Technical Proficiency criterion. This table indicates that 88M was the only MOS with fewer than 100 soldiers conforming to this selection strategy, regardless of the predictor set. Similar results and sample sizes were found for the other criterion constructs (the sample sizes for the full set of performance factor criteria are reported in Appendix B).

Based on these findings, all analyses reported in this chapter were conducted using samples selected with the predictor/criterion setwise deletion strategy. However, because the number of soldiers in 88M was considerably smaller than that of the other MOS, that MOS (along with 31C) was not included in the analyses.

Table 3.3

Soldiers in LVII Sample Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Core Technical Proficiency by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	333	322	112	301	297	309
13B	170	165	152	159	148	156
19K	156	130	130	122	123	129
63B	169	147	139	136	132	140
71L	147	115	104	105	109	102
88M[a]	84	56	54	51	52	53
91A	205	191	174	185	165	183
95B	160	149	140	142	133	145
Total	1,424	1,275	1,005	1,201	1,159	1,217

[a]MOS 88M was not included in LVII validity analyses.

Predictors

The predictor scores used in these analyses were derived from the operationally administered ASVAB and the paper-and-pencil and computerized tests administered in the Project A LV Experimental Battery. For the ASVAB, four types of scores were examined. These scores, listed in Table 3.4, include the nine ASVAB subtests (of which the Verbal score is a composite of the Word Knowledge and Paragraph Comprehension subtests), the four ASVAB factor composite scores, the AFQT, and the MOS-appropriate Aptitude Area composite scores.

The scores derived from the LV Experimental Battery are listed in Table 3.5. With one exception, these scores are described by Peterson et al. (1990). The exception concerns the scores derived for the ABLE. Note that three different sets of ABLE scores are listed in Table 3.5. The first set, labeled the ABLE Rational Composites, were derived along with the other LV predictor composites. The other two sets, labeled ABLE-168 Composites and ABLE-114 Composites, were based on results of factor analyses of the ABLE items. ABLE-168 was scored using 168 of the ABLE items, and ABLE-114 was scored using only 114 items. The development of these scores is described by White (1994).

Table 3.4

Four Sets of ASVAB Scores Used in LVII Validity Analyses

ASVAB Subtests
- General Science
- Arithmetic Reasoning
- Verbal (Word Knowledge and Paragraph Comprehension)
- Numerical Operations
- Coding Speed
- Auto/Shop Information
- Mathematical Knowledge
- Mechanical Comprehension
- Electronic Information

ASVAB Factor Composites
- Technical (Auto/Shop, Mechanical Comprehension, Electronics Information)
- Quantitative (Math Knowledge, Arithmetic Reasoning)
- Verbal (Word Knowledge, Paragraph Comprehension, General Science)
- Speed (Coding Speed, Number Operations)

AFQT

Aptitude Area Composites (1 per MOS)
- 11B: CO (Combat)
- 13B: FA (Field Artillery)
- 19K: CO (Combat)
- 63B: MM (Mechanical Maintenance)
- 71L: CL (Clerical)
- 91A: ST (Skilled Technical)
- 95B: ST (Skilled Technical)

Criteria

The second-tour performance measures collected from the LVII sample constitute a subset and extension of the first-tour measures collected from the LVI sample. These measures include:

- o Revised hands-on job sample measures
- o Revised job knowledge tests
- o Army-wide BARS
- o MOS-specific BARS
- o Combat performance prediction ratings
- o Administrative indices of performance
- o Supervisory simulation exercises
- o Situational Judgment Test

Table 3.5

Sets of Experimental Battery Predictor Scores Used in LVII Validity Analyses

Spatial Composite	**ABLE Rational Composites**
Spatial	Achievement Orientation
	Adjustment
Computer Composites	Physical Condition
Psychomotor	Internal Control
Perceptual Speed	Cooperativeness
Perceptual Accuracy	Dependability
Number Speed and Accuracy	Leadership
Basic Speed	
Basic Accuracy	**ABLE-168 Composites**
Short-Term Memory	Locus of Control
Movement Time	Cooperativeness
	Dominance
JOB Composites	Dependability
Autonomy	Physical Condition
High Expectations	Stress Tolerance
Routine	Work Orientation
AVOICE Composites	**ABLE-114 Composites**
Administrative	Locus of Control
Audiovisual Arts	Cooperativeness
Food Service	Dominance
Structural/Machines	Dependability
Protective Services	Physical Condition
Rugged/Outdoors	Stress Tolerance
Social	Work Orientation
Skilled Technical	

These measures generated a set of 23 basic scores that were the basis for the LVII performance modeling analysis reported by Hanson, Campbell, and McKee (in press). Those analyses examined the fit of a preliminary model of second-tour performance that was based on data collected from the Project A Concurrent Validation second-tour (CVII) sample, as well as the fit of various alternative a priori models. The model of second-tour performance emerging from these analyses (labeled the Leadership Factor model) specified six substantive performance factors and three method factors ("written verbal," "ratings," and "simulation exercise"). The three methods factors were defined to be orthogonal to the substantive factors, but the correlations among the substantive factors were not so constrained.

The six substantive factors and the variables that are scored on each are listed in Table 3.6. Of the six substantive second-tour performance factors, three correspond exactly to performance factors in the model of first-tour performance developed using data collected from the Project A Concurrent Validation first-tour (CVI) sample and confirmed using data collected in LVI (Oppler, Childs, & Peterson, 1994). These factors are Core Technical Proficiency (CTP), General Soldiering Proficiency (GSP), and Physical Fitness and Military Bearing (PFB). Consistent with the procedures used for LVI, the GSP factor scores created for soldiers in MOS 11B are

Table 3.6

LVII Performance Factors and the Basic Criterion Scores That Define Them

- **Core Technical Proficiency (CTP)**
  - Hands-On Test - MOS-Specific Tasks
  - Job Knowledge Test - MOS-Specific Tasks

- **General Soldiering Proficiency (GSP)**
  - Hands-On Test - Common Tasks
  - Job Knowledge Test - Common Tasks

- **Achievement and Effort (AE)**
  - Admin. Index - Number of Awards and Certificates
  - Army-Wide BARS Overall Effectiveness Rating Scale
  - Army-Wide BARS Technical Skill/Effort Ratings Factor
  - Average of MOS BARS Rating Scales
  - Average of Combat Prediction Rating Scales

- **Personal Discipline (PD)**
  - Admin. Index - Number of Articles 15 and Flag Actions
  - Army-Wide BARS Personal Discipline Ratings Factor

- **Physical Fitness and Military Bearing (PFB)**
  - Admin. Index - Physical Readiness Score
  - Army-Wide BARS Physical Fitness/Bearing Rating Factor

- **Leadership (LDR)**
  - Admin. Index - Promotion Grade Deviation Score
  - Army-Wide BARS Leading/Supervising Rating Factor
  - Supervisory Simulation - Disciplinary Structure
  - Supervisory Simulation - Disciplinary Communication
  - Supervisory Simulation - Disciplinary Interpersonal Skill
  - Supervisory Simulation - Counseling Diagnosis/Prescription
  - Supervisory Simulation - Counseling Communication/Interpersonal Skills
  - Supervisory Simulation - Training Structure
  - Supervisory Simulation - Training Motivation
  - Situational Judgment Test - Total Score

treated as CTP scores in the validity analyses. (Tasks that are considered "general" to the Army for soldiers in most other MOS are considered central or "core" to soldiers in 11B.)

Two of the second-tour performance factors--Achievement and Effort (AE) and Personal Discipline (PD)--have a somewhat different composition than their first-tour counterparts (Effort and Leadership [ELS] and Maintaining Personal Discipline [MPD], respectively). That is, the second-tour Achievement and Effort factor contains one score (the average of the Combat Performance Prediction Rating Scales) that was not included in the first-tour Effort and Leadership factor and does not include any rating scales reflective of leadership performance. Also, the second-tour Personal Discipline factor is

missing one score (Promotion Rate) that was incorporated in the first-tour version of that factor. The sixth second-tour performance factor--Leadership (LDR)--has no counterpart in the first-tour performance model, although it does include the Promotion Rate score that had previously been included in the first-tour MPD factor and all rating scales reflective of leadership performance.

In addition to the six performance factors, four additional criteria were used in the validation analyses reported in this chapter. Two of these criteria are variations of the Leadership factor. The first variation (LDR2) does not include the Situational Judgment Test (SJT) total score that was included in the original Leadership factor, and the second variation (LDR3) does not include either the SJT or the scores from the Supervisory Simulation exercises. The other two criteria included in the validation analyses are the total scores from the hands-on and the job knowledge tests.

Analyses

The analysis procedure consisted of the following steps.

A. Using the predictor/criterion setwise deletion sample, multiple correlations between each set of predictor scores and the substantive factor scores, the two additional leadership factor scores, and the total scores from the hands-on and job knowledge tests were computed separately by MOS and then averaged.

1) As indicated above, ASVAB was represented by

   a) The nine ASVAB subtest scores
   b) The four ASVAB factor scores
   c) The AFQT
   d) The MOS-appropriate Aptitude Area composite score

2) ABLE was represented by three sets of scores:

   a) The seven rational scales developed for the Experimental Battery
   b) Seven empirical scales developed on the basis of a factor analysis that retained all of the items (and which used 168 of them)
   c) Seven empirical scales developed by using the results of the factor analysis to select the best items to reflect each factor (and which used only 114 items)

Each of the other predictor sets was represented by a single set of scores as described above and in Table 3.5.

Results were computed both with and without correcting for multivariate range restriction (Lord & Novick, 1968). Corrections for range restriction were made using the 9x9 intercorrelation matrix among the subtests in the 1980 Youth Population (Department of Defense, 1982). All results were adjusted for shrinkage using Rozeboom's (1978) Formula 8.

B. Using the predictor/criterion setwise deletion sample, incremental validities for each set of Experimental Battery predictors (e.g., AVOICE composites or computer composites) over the four ASVAB factor composites were computed against the same criteria used to compute the validities in Step A. Once again, the results were computed separately by MOS and then averaged. Also, the results were computed both with and without correcting for range restriction, and were adjusted for shrinkage using the Rozeboom formula.

RESULTS

Multiple Correlations for ASVAB and LVII Experimental Battery Predictors (Based on Predictor/Criterion Deletion Sample)

Multiple Correlations by Predictor Type

Multiple correlations for the four ASVAB factors, the single spatial composite, the eight computer-based predictor scores, the three JOB composite scores, the seven ABLE composite scores, and the eight AVOICE composite scores are reported in Table 3.7. Using the predictor/criterion deletion sample, these results were computed separately by MOS and then averaged. These results have also been adjusted for shrinkage using the Rozeboom formula and corrected for range restriction. Results which have not been corrected for range restriction (but which have been adjusted for shrinkage) are reported in Table 3.8.

The results in Table 3.7 indicate that the four ASVAB factors were the best set of predictors for all of the criterion performance factors (CTP, GSP, AE, PD, PFB, and LDR), the two additional leadership factors (LD2 and LD3), and the Hands-On and Job Knowledge total scores. The highest multiple correlation was between the ASVAB factors and the Job Knowledge-Total score ($R = .74$), while the lowest was with the PD and PFB scores ($R = .15$ and $.16$, respectively).

With the exception of the prediction of PFB with the ABLE composites, the spatial composite was the next best predictor. The pattern of multiple Rs for the spatial composite was highly similar to the ASVAB pattern, with the highest multiple R being with the JK-Total score ($R = .67$) and the lowest with the PD and PFB scores ($R = .15$ and $.13$, respectively). Next in line were the eight computer composites, except for the prediction of AE, where the ABLE composites had slightly higher validities ($R = .09$ vs. $R = .14$), and of PFB, where the computer composite had the smallest R.

The three JOB composites, the seven ABLE composites, and the eight AVOICE composites had different patterns of multiple Rs for the different criterion performance factors. The AVOICE was highest among the three for CTP, GSP, LDR, and the JK-Total score, while JOB was highest for the HO-Total criterion; ABLE was highest for AE, PFB, LDR2, and LDR3.

Table 3.7

Mean of Multiple Correlations Computed Within Job for LVII Samples for ASVAB Factors, Spatial, Computer, JOB, ABLE Composites, and AVOICE: Corrected for Range Restriction

Criterion[a]	No. of MOS[b]	ASVAB Factors [4]	Spatial [1]	Computer [8]	JOB [3]	ABLE Comp. [7]	AVOICE [8]
CTP	7	64 (10)	57 (11)	53 (11)	33 (17)	24 (19)	41 (12)
GSP	6	63 (06)	58 (05)	48 (10)	28 (19)	19 (17)	29 (24)
AE	7	29 (14)	27 (13)	09 (11)	07 (12)	13 (17)	09 (15)
PD	7	15 (12)	15 (10)	12 (12)	03 (05)	06 (10)	06 (10)
PFB	7	16 (11)	13 (06)	03 (06)	07 (08)	17 (15)	09 (13)
LDR	7	63 (14)	55 (08)	49 (13)	34 (21)	34 (20)	35 (24)
LDR2	7	51 (16)	46 (12)	35 (19)	26 (21)	25 (22)	25 (24)
LDR3	7	47 (13)	39 (12)	31 (15)	19 (18)	23 (17)	20 (21)
HO-Total	7	46 (13)	41 (14)	33 (21)	24 (11)	12 (15)	21 (18)
JK-Total	7	74 (05)	67 (03)	58 (06)	37 (16)	29 (17)	44 (14)

Note: Predictor/criterion setwise deletion sample. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] CTP = Core Technical Proficiency; GSP = General Soldiering Proficiency; AE = Achievement and Effort; PD = Personal Discipline; PFB = Physical Fitness/Military Bearing; LDR = Leadership; LDR2 = Leadership minus Situational Judgment Test; LDR3 = Leadership minus Situational Judgment Test and Supervisory Simulation Exercises; HO = Hands-On; JK = Job Knowledge

[b] Number of MOS for which validities were computed.

Table 3.8

Mean of Multiple Correlations Computed Within Job for LVII Samples for ASVAB Factors, Spatial, Computer, JOB, ABLE Composites, and AVOICE: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors [4]	Spatial [1]	Computer [8]	JOB [3]	ABLE Comp. [7]	AVOICE [8]
CTP	7	40 (12)	34 (12)	20 (15)	11 (13)	00 (00)	13 (15)
GSP	6	48 (06)	42 (04)	21 (15)	13 (15)	02 (05)	15 (13)
AE	7	12 (09)	14 (03)	00 (00)	00 (00)	07 (10)	02 (06)
PD	7	02 (03)	09 (06)	03 (06)	02 (04)	05 (10)	03 (06)
PFB	7	09 (12)	08 (06)	00 (00)	05 (08)	15 (15)	06 (11)
LDR	7	38 (12)	33 (05)	09 (14)	12 (15)	11 (14)	12 (15)
LDR2	7	30 (11)	26 (04)	08 (14)	08 (12)	09 (12)	08 (11)
LDR3	7	26 (07)	19 (05)	06 (12)	03 (07)	04 (07)	06 (10)
HO-Total	7	28 (13)	25 (12)	12 (17)	06 (07)	00 (00)	08 (12)
JK-Total	7	56 (08)	47 (06)	24 (13)	18 (15)	02 (05)	19 (16)

Note: Predictor/criterion setwise deletion sample. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

In general, with regard to the ABLE, the highest correlations are with the Leadership factor and the Core Technical Performance factor. Comparatively, the correlations of the ABLE with Achievement and Effort and with Personal Discipline are lower. In large part this reflects the emergence of a separate leadership factor and the fact that the promotion rate index produced a better fit for the LVII model when it was scored as a Leadership component than as a component of the Personal Discipline factor. As in CVII, a faster promotion rate for LVII personnel is more a function of good things that happen rather than an absence of negative events that act to slow an individual's progression, as it was in CVI and LVI.

When these results are not corrected for range restriction (Table 3.8), all of the multiple Rs were lowered, some quite dramatically. Overall, the ASVAB factors still had the highest multiple Rs, with the spatial composite again having a highly similar (but slightly smaller in magnitude) pattern. (It should be noted that differences between the patterns of the corrected and uncorrected results could arise due to differences in range restriction across the various predictors. Therefore, where differences between the two sets of results are reported, we believe that the corrected results should be given greater emphasis, since differences between measures with respect to range restriction will have been removed.)

Comparisons of Alternative ASVAB Scores

The average multiple correlations (corrected and uncorrected for range restrictions, respectively) for the four different sets of ASVAB scores are reported in Tables 3.9 and 3.10. The corrected results indicate that all four sets had virtually identical multiple Rs with all of the criterion measures. The four ASVAB factors tended to have very slightly higher validities than the other three sets of scores, whereas the AFQT tended to have slightly lower validities.

The uncorrected correlations again showed highly consistent multiple Rs for the four different sets of ASVAB scores (Table 3.10); as was true previously, the uncorrected Rs were much smaller than the corrected correlations. For a few criteria -- AE, PFB, and LDR2 and LDR3 -- the multiple Rs for ASVAB subtests were smaller than the other three sets.

Comparisons of Alternative ABLE Scores

The average multiple correlations for the three sets of ABLE scores are reported in Tables 3.11 and 3.12 (corrected and uncorrected, respectively). The results in Table 3.11 indicate that the pattern and levels of multiple Rs were generally very similar across the three sets. However, the ABLE composites were somewhat better predictors of LDR ($R = .34$) than were the ABLE-168 and ABLE-114 factor scores ($R = .26$ and $.27$, respectively), and the ABLE-168 factor scores were the best predictors of CTP ($R = .30$ versus $R = .24$ for the ABLE composites and $R = .27$ for the ABLE-114 factor scores). Note that the uncorrected results reported in Table 3.11 are much smaller than the results in Table 3.10, but the pattern and levels of the validities are still very similar across the three sets of scores.

Table 3.9

Mean of Multiple Correlations Computed Within Job for LVII Samples for ASVAB Subtests, ASVAB Factors, AFQT, and ASVAB MOS-Appropriate Aptitude Area: Corrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Subtests [9]	ASVAB Factors [4]	AFQT [1]	Aptitude Area[b] [1]
CTP	7	64 (10)	64 (10)	61 (08)	63 (09)
GSP	6	62 (07)	63 (06)	58 (09)	61 (07)
AE	7	28 (10)	29 (14)	28 (13)	30 (11)
PD	7	12 (11)	15 (12)	16 (11)	16 (10)
PFB	7	10 (08)	16 (11)	16 (06)	14 (07)
LDR	7	64 (10)	63 (14)	62 (15)	59 (14)
LDR2	7	51 (13)	51 (16)	50 (16)	49 (16)
LDR3	7	46 (13)	47 (13)	45 (14)	44 (14)
HO-Total	7	45 (12)	46 (13)	43 (13)	46 (11)
JK-Total	7	74 (05)	74 (05)	70 (06)	72 (06)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.
[b] MOS-Appropriate Aptitude Area composite.

Table 3.10

Mean of Multiple Correlations Computed Within Job for LVII Samples for ASVAB Subtests, ASVAB Factors, AFQT, and ASVAB MOS-Appropriate Aptitude Area: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Subtests [9]	ASVAB Factors [4]	AFQT [1]	Aptitude Area[b] [1]
CTP	7	40 (10)	40 (12)	39 (06)	42 (09)
GSP	6	45 (08)	48 (06)	43 (08)	44 (06)
AE	7	03 (07)	12 (09)	14 (08)	16 (08)
PD	7	01 (02)	02 (03)	11 (06)	09 (06)
PFB	7	02 (05)	09 (12)	10 (09)	09 (08)
LDR	7	36 (17)	38 (12)	40 (10)	35 (09)
LDR2	7	25 (14)	30 (11)	30 (10)	28 (10)
LDR3	7	19 (13)	26 (07)	24 (11)	24 (11)
HO-Total	7	24 (14)	28 (13)	27 (08)	31 (08)
JK-Total	7	55 (09)	56 (08)	51 (05)	52 (06)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.
[b] MOS-Appropriate Aptitude Area composite.

Table 3.11

Mean of Multiple Correlations Computed Within Job for LVII Samples for ABLE Composites, ABLE-168, and ABLE-114: Corrected for Range Restriction

Criterion	No. of MOS[a]	ABLE Composites [7]	ABLE-168 [7]	ABLE-114 [7]
CTP	7	24 (19)	30 (10)	27 (15)
GSP	6	19 (17)	18 (16)	18 (17)
AE	7	13 (17)	12 (14)	11 (17)
PD	7	06 (10)	05 (09)	04 (07)
PFB	7	17 (15)	16 (16)	16 (14)
LDR	7	34 (20)	26 (25)	27 (25)
LDR2	7	25 (22)	24 (23)	25 (23)
LDR3	7	23 (17)	19 (18)	20 (19)
HO-Total	7	12 (15)	12 (13)	14 (14)
JK-Total	7	29 (17)	30 (13)	28 (15)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.12

Mean of Multiple Correlations Computed Within Job for LVII Samples for ABLE Composites, ABLE-168, and ABLE-114: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ABLE Composites [7]	ABLE-168 [7]	ABLE-114 [7]
CTP	7	00 (00)	00 (00)	00 (00)
GSP	6	02 (05)	03 (06)	00 (00)
AE	7	07 (10)	03 (05)	04 (08)
PD	7	05 (10)	05 (10)	04 (10)
PFB	7	15 (15)	15 (14)	16 (13)
LDR	7	11 (14)	09 (10)	08 (11)
LDR2	7	09 (12)	08 (10)	07 (08)
LDR3	7	04 (07)	00 (00)	00 (00)
HO-Total	7	00 (00)	01 (03)	01 (04)
JK-Total	7	02 (05)	03 (07)	00 (00)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Incremental Validities for the Experimental Battery Predictors Over the ASVAB Factors (Based on Predictor/Criterion Deletion Sample)

Incremental validity results for the Experimental Battery predictors over the ASVAB factor composites are reported in Tables 3.13 through 3.16. (Results corrected for range restriction are reported in Tables 3.13 and 3.14, uncorrected results in Tables 3.15 and 3.16.) Table 3.13 reports the multiple correlations for the four ASVAB factor composites alone (as computed separately in each of the predictor/criterion deletion samples), whereas Table 3.14 reports the multiple correlations for the four ASVAB factors along with each set of predictors in the Experimental Battery. Numbers underlined in Table 3.14 indicate multiple correlations that are higher than those based on ASVAB alone.

The results indicate that there were no increments to the prediction of any of the criteria for the computer, JOB, or AVOICE composites. The spatial composite added a point to the prediction of GSP, AE, and JK-Total, and the ABLE composites added five points to the prediction of PFB (from R = .13 for ASVAB alone to R = .18 for ASVAB and ABLE together) and one point to the prediction of LDR (from R = .64 to R = .65).

The incremental validity results that are not corrected for range restriction (Tables 3.15 and 3.16) are similar in pattern to the corrected results. Again, the largest increase in validity was for the PFB criterion factor when the ABLE composites were added to the ASVAB factor composites -- a change from R = .10 to R = .17.

The adjusted incremental validity results for each of the three sets of ABLE scores over the ASVAB factors are reported in Tables 3.17 and 3.18 (corrected and uncorrected for range restriction, respectively). The corrected results in Table 3.17 indicate that all three sets of scores account for variance in PFB not predicted by the ASVAB factors, and the ABLE composites and ABLE-168 factor scores add to the prediction of LDR. In comparison, the uncorrected results indicate that all three sets of scores provided incremental validity for both PD and PFB, but only the ABLE composites add to the prediction of LDR.

Table 3.13

Mean of Multiple Correlations Computed Within Job for ASVAB Factors Within Each LVII Predictor/Criterion Setwise Deletion Sample: Corrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (Spatial Sample) [4]	ASVAB Factors (Computer Sample) [4]	ASVAB Factors (JOB Sample) [4]	ASVAB Factors (ABLE Comp. Sample) [4]	ASVAB Factors (AVOICE Sample) [4]
CTP	7	65 (10)	64 (07)	66 (11)	67 (09)	66 (12)
GSP	6	61 (08)	62 (09)	60 (09)	61 (11)	61 (08)
AE	7	31 (15)	24 (13)	32 (15)	30 (17)	30 (16)
PD	7	17 (12)	21 (17)	17 (13)	15 (11)	16 (12)
PFB	7	15 (10)	17 (14)	16 (10)	13 (13)	18 (11)
LDR	7	63 (13)	62 (16)	63 (14)	64 (12)	62 (13)
LDR2	7	52 (17)	50 (20)	50 (19)	51 (20)	49 (19)
LDR3	7	46 (19)	45 (23)	44 (22)	45 (23)	43 (23)
HO-Total	7	46 (14)	45 (17)	44 (17)	46 (18)	47 (16)
JK-Total	7	74 (05)	74 (06)	74 (06)	75 (05)	74 (06)

Note: Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.14

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVII Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE: Corrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (A4) + Spatial [5]	A4 + Computer [12]	A4 + JOB [7]	A4 + ABLE Composites [11]	A4 + AVOICE [12]
CTP	7	65 (11)	63 (10)	65 (11)	65 (11)	64 (13)
GSP	6	**62** (08)	62 (10)	60 (11)	60 (10)	57 (12)
AE	7	**33** (10)	10 (14)	30 (15)	24 (20)	20 (18)
PD	7	16 (12)	16 (17)	14 (15)	13 (12)	06 (11)
PFB	7	12 (10)	08 (11)	13 (11)	**18** (15)	11 (16)
LDR	7	63 (12)	61 (15)	63 (13)	**65** (13)	62 (13)
LDR2	7	51 (17)	48 (20)	49 (20)	50 (24)	48 (19)
LDR3	7	45 (20)	41 (21)	42 (23)	45 (22)	40 (24)
HO-Total	7	46 (15)	43 (17)	44 (15)	43 (21)	39 (23)
JK-Total	7	**75** (05)	73 (06)	74 (06)	74 (06)	73 (06)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone (as reported in Table 3.13). Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.15

Mean of Multiple Correlations Computed Within Job for ASVAB Factors Within LVII Predictor/Criterion Setwise Deletion Sample: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (Spatial) [4]	ASVAB Factors (Computer) [4]	ASVAB Factors (JOB) [4]	ASVAB Factors (ABLE Comp.) [4]	ASVAB Factors (AVOICE) [4]
CTP	7	41 (11)	41 (10)	42 (11)	42 (09)	42 (11)
GSP	6	45 (09)	46 (09)	44 (11)	44 (11)	44 (12)
AE	7	13 (10)	07 (07)	16 (08)	12 (12)	12 (08)
PD	7	04 (06)	05 (06)	04 (07)	02 (04)	03 (04)
PFB	7	08 (11)	09 (13)	09 (12)	10 (13)	08 (11)
LDR	7	39 (09)	40 (10)	38 (08)	38 (11)	37 (10)
LDR2	7	29 (12)	29 (14)	28 (15)	27 (15)	27 (15)
LDR3	7	25 (12)	26 (12)	25 (13)	25 (12)	25 (12)
HO-Total	7	26 (16)	28 (13)	24 (18)	25 (15)	27 (15)
JK-Total	7	56 (08)	55 (10)	55 (07)	56 (08)	54 (09)

Note: Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.16

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVII Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (A4) + Spatial [5]	A4 + Computer [12]	A4 + JOB [7]	A4 + ABLE Composites [11]	A4 + AVOICE [12]
CTP	7	41 (13)	37 (15)	40 (12)	35 (14)	36 (16)
GSP	6	<u>47</u> (09)	45 (10)	42 (16)	42 (10)	38 (19)
AE	7	<u>14</u> (07)	00 (00)	09 (09)	10 (11)	03 (09)
PD	7	03 (05)	03 (08)	<u>06</u> (10)	<u>05</u> (09)	01 (03)
PFB	7	07 (10)	03 (05)	08 (11)	<u>17</u> (13)	08 (13)
LDR	7	39 (09)	31 (18)	37 (11)	<u>39</u> (13)	34 (17)
LDR2	7	28 (12)	20 (20)	27 (15)	26 (17)	22 (17)
LDR3	7	23 (12)	14 (13)	20 (15)	19 (15)	17 (14)
HO-Total	7	26 (17)	22 (20)	22 (17)	21 (16)	18 (19)
JK-Total	7	<u>57</u> (08)	53 (10)	55 (09)	53 (10)	50 (14)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone (as reported in Table 3.15). Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.17

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVII Samples for ABLE Composites, ABLE-168, and ABLE-114: Corrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (A4) [4]	A4 + ABLE Comp. [11]	A4 + ABLE-168 [11]	A4 + ABLE-114 [11]
CTP	7	*67* (09)	65 (11)	65 (11)	65 (11)
GSP	6	*61* (11)	60 (10)	59 (13)	59 (13)
AE	7	*30* (17)	24 (20)	19 (22)	21 (21)
PD	7	*15* (11)	13 (12)	14 (13)	13 (11)
PFB	7	*13* (13)	**18** (15)	**16** (16)	**17** (17)
LDR	7	*64* (12)	**65** (13)	**65** (13)	64 (13)
LDR2	7	*51* (20)	50 (24)	50 (25)	49 (25)
LDR3	7	*45* (23)	45 (22)	45 (22)	44 (22)
HO-Total	7	*46* (18)	43 (21)	42 (21)	42 (21)
JK-Total	7	*75* (05)	74 (06)	74 (06)	74 (06)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Multiple Rs for ASVAB Factors alone are in italics. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.18

Mean of Incremental Correlations Over ASVAB Factors Computed Within Job for LVII Samples for ABLE Composites, ABLE-168, and ABLE-114: Uncorrected for Range Restriction

Criterion	No. of MOS[a]	ASVAB Factors (A4) [4]	A4 + ABLE Comp. [11]	A4 + ABLE-168 [11]	A4 + ABLE-114 [11]
CTP	7	*42* (09)	35 (14)	36 (13)	35 (12)
GSP	6	*44* (11)	42 (10)	39 (20)	38 (20)
AE	7	*12* (12)	10 (11)	05 (08)	08 (10)
PD	7	*02* (04)	**05** (09)	**06** (11)	**05** (09)
PFB	7	*10* (13)	**17** (13)	**15** (14)	**16** (15)
LDR	7	*38* (11)	**39** (13)	**38** (13)	**38** (12)
LDR2	7	*27* (15)	26 (17)	27 (16)	26 (16)
LDR3	7	*25* (12)	19 (15)	20 (13)	17 (13)
HO-Total	7	*25* (15)	21 (16)	17 (18)	18 (18)
JK-Total	7	*56* (08)	53 (10)	53 (11)	53 (10)

Note: Predictor/criterion setwise deletion samples. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Numbers in parentheses are standard deviations. Multiple Rs for ASVAB Factors alone are in italics. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone. Decimals omitted.

[a] Number of MOS for which validities were computed.

COMPARISON BETWEEN VALIDITY RESULTS OBTAINED WITH LVI AND LVII SAMPLES

The final set of results concern the comparison between the validity results computed for the first-tour Longitudinal Validation (LVI) sample and those reported for the second-tour Longitudinal Validation (LVII) sample.

Levels of Validity

The multiple correlations for the ASVAB factors and each set of experimental predictors as computed for LVI and LVII, respectively, are reported in Table 3.19. These results have been corrected for range restriction and adjusted for shrinkage.

Note that the first-tour results are based on the LVI samples selected using the setwise deletion strategy described above for LVI, in comparison to the predictor/criterion setwise deletion strategy used for LVII. Thus, the soldiers included in the LVI validation samples were required to have complete criterion data, but the soldiers in the LVII sample were not. Also, the LVII analyses did not include two MOS (31C and 88M) that were included in the LVI analyses. Finally, as described earlier, there were differences between the components of the Achievement and Effort (AE) and Personal Discipline (PD) factors computed for soldiers in the LVII sample and their corresponding factors in the LVI sample (Effort and Leadership [ELS] and Maintaining Personal Discipline [MPD], respectively).

The results in Table 3.19 demonstrate that the patterns and levels of validities are very similar across the two sets of analyses, especially for the four ASVAB factor composites. For example, the multiple R between the ASVAB and the CTP is .63 and .64 for the LVI and LVII samples, respectively.

The greatest discrepancies between the two sets of results concern the multiple correlations between the ABLE composites and two of the three "will do" criterion factors--[Maintaining] Personal Discipline and Physical Fitness and Military Bearing. Specifically, the multiple R between ABLE and the discipline factor decreases from .15 in LVI to .06 in LVII, and the multiple R between ABLE and Physical Fitness and Military Bearing decreases from .28 to .17.

Some of the decrease in the ABLE's ability to predict Personal Discipline in LVII may be due to the removal from that factor of the Promotion Rate component. The validities of the other predictors were not similarly affected by this scoring change. Again, the highest correlation for the ABLE in LVII was with the Leadership factor ($R = .34$). The LVII Leadership factor included the promotion rate index, all scores derived from the supervisory role plays, <u>and</u> the Army-wide BARS (rating scale) leading and supervising factor, which was part of the ELS factor in CVI and LVI. In effect, these differences were expected to decrease the ABLE correlations with the LVII Achievement and Effort factor and increase the ASVAB correlations with this factor, which in LVII is more reflective of technical achievement than was the ELS factor in LVI and CVI. These are the expected patterns and they lend further support to the construct validity of the performance models.

Table 3.19

Comparison of Mean Multiple Correlations Computed Within Job for ASVAB Factors, Spatial, Computer, JOB, ABLE Composites, and AVOICE Within LVI and LVII Samples: Corrected for Range Restriction

	No. of MOS[b]		ASVAB Factors [4]		Spatial [1]		Computer [8]		JOB [3]		ABLE Comp. [7]		AVOICE [8]	
Criterion[a]	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII
CTP	9	7	63	64	58	57	49	53	31	33	21	24	39	41
GSP	8	6	66	63	65	58	55	48	32	28	24	19	38	29
ELS/AE	9	7	34	29	33	27	30	09	19	07	12	13	20	09
MPD/PD	9	7	16	15	14	15	10	12	06	03	15	06	05	06
PFB	9	7	12	16	08	13	13	03	07	07	28	17	09	09
HO-Total	9	7	50	46	50	41	38	33	20	24	13	12	30	21
JK-Total	9	7	73	74	66	67	60	58	38	37	30	29	43	44

Note: LVI setwise deletion samples; LVII predictor/criterion setwise deletion sample. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Decimals omitted.

[a] CTP = Core Technical Proficiency; GSP = General Soldiering Proficiency; ELS = Effort and Leadership (LVI); AE = Achievement and Effort (LVII); MPD = Maintaining Personal Discipline (LVI); PD = Personal Discipline (LVII); PFB = Physical Fitness/Military Bearing; HO = Hands-On; JK = Job Knowledge
[b] Number of MOS for which validities were computed.

Incremental Validities

Finally, the incremental validity results for the LVI and LVII analyses are reported in Tables 3.20 and 3.21. These results are also corrected for range restriction and adjusted for shrinkage. Table 3.20 reports the multiple correlations for the four ASVAB factor composites alone (as computed separately in each of the LVI setwise deletion samples and each of the LVII predictor/criterion setwise deletion samples), whereas Table 3.21 reports the multiple correlations for the four ASVAB factors along with each set of predictors in the Experimental Battery. Numbers underlined in Table 3.21 indicate multiple correlations that are higher than those based on ASVAB alone.

Generally, the results are very similar across the two samples. For example, the results indicate that the spatial composite contributed a small amount of incremental validity in both the LVI and LVII samples, and that most of the other predictors added less. Once again, the primary discrepancy between the two sets of results concerns the ABLE composites. That is, whereas the ABLE composites added eight points to the prediction of MPD (from R = .16 for ASVAB alone to R = .24 for ASVAB and ABLE together) and 17 points to the prediction of PFB (from R = .15 to R = .32) in LVI, they added nothing to the prediction of PD in LVII and only five points to the prediction of PFB (from R = .13 to R = .18).

Table 3.20

Comparison of Mean Multiple Correlations Computed Within Job for ASVAB Factors Within LVI Setwise Deletion and LVII Predictor/Criterion Setwise Deletion Samples: Corrected for Range Restriction

Criterion	No. of MOS[a]		ASVAB Factors (Spatial) [5]		ASVAB Factors (Computer) [12]		ASVAB Factors (JOB) [7]		ASVAB Factors (ABLE Comp.) [11]		ASVAB Factors (AVOICE) [12]	
	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII
CTP	9	7	63	65	62	64	63	66	62	67	64	66
GSP	8	6	66	61	65	62	67	60	66	61	67	61
ELS/AE	9	7	37	31	37	24	37	32	36	30	37	30
MPD/PD	9	7	16	17	15	21	15	17	16	15	16	16
PFB	9	7	16	15	19	17	16	16	15	13	16	18
HO-Total	9	7	51	46	50	45	50	44	50	46	51	47
JK-Total	9	7	71	74	71	74	72	74	71	75	72	74

Note: Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Decimals omitted.

[a] Number of MOS for which validities were computed.

Table 3.21

Comparison of Mean Incremental Correlations Over ASVAB Factors Computed Within Job for LVI and LVII Samples for Spatial, Computer, JOB, ABLE Composites, and AVOICE: Corrected for Range Restriction

Criterion	No. of MOS[a]		ASVAB Factors (A4) + Spatial [5]		A4 + Computer [12]		A4 + JOB [7]		A4 + ABLE Comp. [11]		A4 + AVOICE [12]	
	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII	LVI	LVII
CTP	9	7	64	65	61	63	63	65	61	65	64	64
GSP	8	6	69	62	66	62	67	60	66	60	66	57
ELS/AE	9	7	37	33	36	10	37	30	36	24	36	20
MPD/PD	9	7	15	16	15	16	12	14	24	13	11	06
PFB	9	7	15	12	17	08	17	13	32	18	15	11
HO-Total	9	7	53	46	49	43	50	44	49	43	50	39
JK-Total	9	7	73	75	71	73	72	74	71	74	71	73

Note: LVI setwise deletion sample; LVII predictor/criterion setwise deletion sample. Adjusted for shrinkage (Rozeboom formula 8). Numbers in brackets are the numbers of predictor scores entering prediction equations. Underlined numbers denote multiple Rs greater than for ASVAB Factors alone (as reported in Table 3.20). Decimals omitted.

[a] Number of MOS for which validities were computed.

SUMMARY AND CONCLUSION

The preceding analyses of basic validation results for the LVII sample produced results that were largely consistent with those obtained during the basic validation analyses for LVI. In summarizing the prior validation results, Oppler, Peterson, and Russell (1994) concluded that the ASVAB was the best overall predictor of first-tour performance, but that the composite of spatial tests provided a small amount of incremental validity for the "can do" criteria (i.e., Core Technical Proficiency, General Soldiering Proficiency), and the ABLE provided larger increments for two of the three "will do" criteria (Maintaining Personal Discipline, and Physical Fitness and Military Bearing). The same pattern of results was found in the present analyses.

Furthermore, not only were the LVII results similar in pattern to those of the LVI analyses, they were also similar in magnitude (with the notable exception of some of the ABLE validities, which were substantially lower in LVII). In particular, the multiple correlations between the ASVAB factor composites and second-tour criterion scores were rarely more than three points lower than the multiple correlations between the ASVAB factors and the corresponding first-tour criteria. Given the length of time between the collection of the first- and second-tour performance measures (approximately three years), this level of decrement in validities is remarkably small.

With regard to the ASVAB and options in using subtest scores to form prediction equations, the LVII results indicate highly similar results across the four methods examined (i.e., nine subtests vs. four factor composites vs. AFQT vs. MOS-specific Aptitude Area composites), with a slight advantage going to the equations using the four factors. Such results were also reported for the LVI sample.

Contrary to results reported in LVI, however, were the relative validities of the three ABLE scoring options examined. In LVI, the method using factor scores computed from the subset of 114 ABLE items proved to have generally higher validities, although the differences were not large. In comparison, the LVII results indicate that the validities of the three sets of scores were very similar, although those associated with the ABLE composites tended to be somewhat greater than those associated with the two sets of factor scores.

Finally, the results of the present analyses indicate that the ASVAB is an extremely valid predictor of the Leadership factor (R = .63 for the four ASVAB factor composites), as are the other cognitive predictors in the Experimental Battery. In fact, none of the predictor sets (including the JOB, ABLE composites, and AVOICE) had multiple correlations of less than .34 with this criterion. Furthermore, multiple correlations between ASVAB and the two modified versions of this score (R = .51 for LDR2, and R = .47 for LDR3) indicate that the relationships between the ASVAB and the Leadership factor cannot be explained purely as being due to shared "written verbal" method variance: The only "paper-and-pencil" component of the Leadership factor--the Situational Judgment Test--was not used in computing LDR2 or LDR3.

Chapter 4
THE PREDICTION OF FUTURE PERFORMANCE FROM CURRENT PERFORMANCE AND FROM TRAINING PERFORMANCE

John P. Campbell, Norman G. Peterson, and Jeffrey Johnson

The general question of how accurately individual job performance at one level in the organization predicts job performance at another level is virtually a "classic problem" in personnel research. It has critical implications for personnel management as well. That is, to what extent should selection for a different job or promotion to a higher level position be based on an evaluation of an individual's performance in the previous job, as compared to alternative types of information that might be relevant? In the Army context, it is a question of the extent to which promotion or reenlistment decisions should be based on assessments of prior performance. A related issue is the extent to which initial job assignments should be based on assessments of performance in training.

TWO GENERAL ISSUES

This general question encompasses a number of more specific queries about the nature of performance in specific organizational contexts. Two of these questions follow. First, the degree to which individual differences in future performance can be predicted from individual differences in past performance is a function of the relative stability of performance across time. Do the true scores for individuals change at different rates even when all individuals are operating under the same "treatment" conditions? The recent arguments over this question sometimes become a bit heated (Ackerman, 1989; Austin, Humphreys, & Hulin, 1989; Barrett, Caldwell, & Alexander, 1985; Barrett & Alexander, 1989; Barrett, Alexander, & Doverspike, 1992; Henry & Hulin, 1987; Hulin, Henry, & Noon, 1990).

The second question concerns whether the current and future jobs possess enough communality in their knowledge, skill, and ability requirements to produce valid predictions of future performance from past performance. Perhaps the new job is simply too different from the old one. For example, the degree to which "managers" should possess domain-specific expertise has long been argued. Just as an army should not be equipped and trained to fight only the last war, the promotion system should not try to maximize performance in the <u>previous</u> job. One implication of this issue depends on whether similar and dissimilar components of performance for the two jobs can be identified and measured. If they can, then the pattern of correlations across performance components can be predicted, and the predictions evaluated.

The data from Project A/Career Force (Campbell & Zook, 1991, 1990) permit some of the above issues to be addressed. Extensive job analyses, criterion development, and analyses of the latent structure of MOS performance for both first tour and second tour have attempted to produce a complete specification of performance at each level. The models of performance for training performance, first-tour performance, and second-tour performance that are summarized in Chapter 1 provide some clear predictions about the pattern of convergent and divergent relationships that should be found. The results

of the prior job analyses also suggest that while the new NCO (second tour) is beginning to acquire supervisory responsibilities, there is a great deal of communality across levels in terms of technical task responsibilities. Consequently, first- and second-tour performance should have a substantial proportion of common determinants.

The three samples of interest are from the longitudinal cohort that entered the Army in 1986-87. This group took the Experimental Test Battery at the start of basic training. Their performance in training was assessed at the end of their technical training course and their job performance was assessed approximately 12-24 months after they entered the Army. For those who reenlisted, their second-tour performance as a junior NCO was assessed during the second half of their second tour. Sample sizes for this latter group were reduced somewhat when soldiers became unavailable for testing during Operation Desert Shield/Desert Storm.

SPECIFIC OBJECTIVES

•Within this context, the specific questions addressed are the following:

(1) To what degree does an individual's level of performance in a first-tour enlisted position predict performance as a junior NCO during his or her second tour?

(2) To what degree does performance in training predict subsequent job performance, both in first-tour enlisted positions and in second-tour positions?

(3) Given that performance is not unidimensional, do the separately measured components of performance exhibit the appropriate patterns of convergent and divergent validity when current performance is used to predict future performance?

(4) As a predictor of future performance, do measures of current performance add variance that is not accounted for by measures of ability, personality, and interests?

METHOD

The samples, predictor tests, and performance measures that are used in this analysis have been described in detail elsewhere (Campbell & Zook, 1991). A brief recapitulation follows.

Samples

The results reported in this chapter are based on the three follow-ups of enlisted personnel in the nine Batch A MOS in the Longitudinal Validation sample who joined the Army in 1986-87.

The sampling plan itself is described in Campbell and Zook (1991). The nine MOS are intended to be representative of Army enlisted positions and the original goal was to begin with a large enough number of new accessions in each MOS to yield 500 soldiers per MOS midway through the first tour and 150 soldiers per MOS midway through the second tour. Because of differential attrition and reenlistment rates, the logistical consequences resulting from the Persian Gulf conflict, and constraints on the number of data collection sites that could be visited, the original goals were not reached in every case. The totals for each MOS at each measurement point are shown in Table 4.1. These are the maximum possible sample sizes. Because of missing data considerations, the Ns for any specific analysis will be smaller.

Table 4.1

Sample Sizes From Each Batch A MOS When Performance was Assessed at Three Points in Time for the Project A/Career Force LV Sample

Batch A MOS		Sample Size (Before Data Editing)		
		End-of-Training (EOT) Performance	First-Tour (LVI) Job Performance	Second-Tour (LVII) Job Performance
11B	Infantryman	8,117	909	47
13B	Cannon Crewmember	4,712	916	80
19E/K	Armor Crewman	2,048	1,073	168
31C	Single Channel Radio Operator	667	529	70
63B	Light-Wheel Vehicle Mechanic	1,215	752	194
71L	Administrative Specialist	1,414	678	157
88M	Motor Transport Operator	1,354	682	89
91A/B	Medical Specialist	3,218	824	222
95B	Military Police	3,639	452	168
	Total	26,384	6,815	1,595

The Longitudinal Predictor Sample

The LV sample consisted of approximately 45,000 new accessions who took the Experimental Predictor Battery during their first three days in the Army. Of the total sample, approximately 30,000 would enter the nine Batch A MOS. The tests that make up the Experimental Predictor Battery and the basic predictor scores that were derived from them were described in Chapter 1.

The End-of-Training (EOT) Sample

The EOT sample for the Batch A MOS consists of those individuals who completed their Advanced Individual Training (AIT) and from whom a set of End-of-Training performance measures were obtained during the last two days of the training period. The measures consisted of the project-developed End-of-

Training achievement test and a set of multiple peer ratings on ten rating scales dealing with technical competence, personnel discipline, effort level, leadership potential, and military bearing. Confirmatory factor analysis techniques were used to develop composites of the individual measures to reflect scores on six latent factors of performance in training. The six performance factors and the individual measures that represent them are shown in Figure 4.1. The development of this factor structure is described in McCloy and Oppler (1990) and summarized in Chapter 1 of this report.

SCORES BASED ON RATING SCALES (PEER RATINGS)

1) Effort and Technical Skill (ETS) Score

    Degree of effective acquisition of technical knowledge/skill

    Degree to which individual demonstrates extra effort

2) Maintaining Personal Discipline (MPD) Score

    Degree to which individual adheres to regulations and orders

    Degree to which individual practices effective self-control

3) Physical Fitness and Military Bearing (PFB)

    Degree to which individual maintains proper military appearance

    Degree to which individual maintains military standards of physical fitness

4) Leadership Potential (LEAD)

    Degree of expected leadership effectiveness

SCORES BASED ON TRAINING ACHIEVEMENT TESTS

5) Basic Knowledge Score

    Items measuring knowledge requirements common to all MOS

6) Technical Knowledge Score

    Items measuring technical knowledge requirements specific to each MOS

Figure 4.1. Six scores based on measures of training performance obtained at the end of basic and technical training.

The Longitudinal Validation First-Tour Performance Assessment Sample (LVI)

The LVI performance measurement sample consisted of all the individuals in the original LV predictor sample who were available for performance assessment 12-24 months later at any one of 15 data collection sites in the United States, Europe, and Korea. For the Batch A MOS this was approximately 6,800 individuals.

Each person underwent one full day of performance assessment. The measures included a hands-on job sample test, a paper-and-pencil test of job knowledge, a number of administrative/archival performance indices, and a number of individual rating scales. These first-tour job performance measures are described in detail in Campbell and Zook (1991) and summarized in Chapter 1 of this report. The individual measures were subjected to a confirmatory modeling analysis (Oppler et al., 1994) and five simple sum factor scores were subsequently used as the best representation of the latent structure of job performance during the first tour of duty. The five performance factors and the individual measures that define them are shown in Figure 4.2.

---

1) Core Technical Proficiency (CTP)
   - Hands-on Test - MOS-Specific Tasks
   - Job Knowledge Test - MOS-Specific Tasks

2) General Soldiering Proficiency (GSP)
   - Hands-On Test - Common Tasks
   - Job Knowledge Test - Common Tasks

3) Effort and Leadership (ELS)
   - Admin. Index - Number of Awards and Certificates
   - Army-Wide BARS Overall Effectiveness Rating Scale
   - Army-Wide BARS Effort/Leadership Ratings Factor
   - Average of MOS BARS Ratings Scales

4) Maintaining Personal Discipline (MPD)
   - Admin. Index - Number of Articles 15 and Flag Actions
   - Admin. Index - Promotion Grade Deviation Score
   - Army-Wide BARS Personal Discipline Ratings Factor

5) Physical Fitness and Military Bearing (PFB)
   - Admin. Index - Physical Readiness Score
   - Army-Wide BARS Fitness/Bearing Ratings Factor

Figure 4.2. LVI first-tour performance factors and the basic criterion scores that define them.

The Longitudinal Validation Second-Tour Performance Measurement Sample (LVII)

The final longitudinal follow-up of the 1986-87 cohort focused on the soldiers in the Batch A MOS who reenlisted for a second tour and who could be located at any one of 16 data collection sites at the time of the LVII data collection. The total sample size across the nine MOS turned out to be approximately 1,550 individuals.

Each individual was assessed for one full day on the set of performance measures developed for assessing second-tour performance. The development of these measures is described in detail in Campbell (1991). As with the previous measurement of the first-tour performance sample, the individual performance measures consisted of a hands-on job sample test, a comprehensive job knowledge test, administrative/archival indices of performance, and multiple rating scales. In addition, the assessment measures for second-tour performance included a paper-and-pencil test of situational leadership/ supervisory judgment and three role-play simulations designed to measure certain aspects of supervisory/leadership skill having to do with counseling and training subordinates. The LVII array of second-tour performance measures is summarized in Chapter 1 of this report.

As described by Hanson et al. (in press), the second-tour performance measures were subjected to a confirmatory analysis and a six-factor solution represented the best fit for both the LVII and the CVII sample data. The six performance factors and the measures that define them are shown in Figure 4.3.

Summary

To summarize, after the initial administration of the Experimental Predictor Battery, the sample of soldiers from the nine Batch A MOS in the 1986-87 cohort was assessed on performance (a) at the end of training, (b) toward the end of their first tour of duty, and (c) for those who reenlisted, toward the end of their second tour of duty. All performance indicators were based on extensive job analyses, used multiple measurement methods, and attempted to represent the content of each job, or training experience, as fully as possible.

Extensive confirmatory analysis at each organizational level yielded a very consistent picture of the latent structure of performance, both across cohorts within levels (i.e., CVI and LVI, and CVII and LVII) and across levels within cohorts (e.g., EOT vs. LVI vs. LVII). The substantive content, or latent structure, of performance showed a strong tendency to be consistent where it should be (i.e., across cohorts) and different where it should be (i.e., across organizational levels).

Latent Variable	Scores Loading on Latent Variables
Core Technical Proficiency (CTP)	Job-Specific Hands-On score Job-Specific Knowledge test score
General Soldiering Proficiency (GSP)	General Hands-On score General Job Knowledge test score
Achievement and Effort (AE)	Adm: Awards Army-Wide BARS ratings: Technical Skill/Effort score Overall Effectiveness rating MOS BARS ratings: Average across all ratings
Maintaining Personal Discipline (MPD)	Adm: Disciplinary Actions (reversed score) Army-Wide BARS ratings: Discipline rating score
Physical Fitness and Military Bearing (PFB)	Adm: Physical Readiness score Army-Wide BARS ratings: Fitness/Bearing score
Leadership (LEAD)	Adm: Promotion Rate Army-Wide BARS ratings: Leading/Supervisory score Discipline Role Play: Structure score Discipline Role Play: Communication score Discipline Role Play: Interpersonal Skill score Counseling Role Play: Diagnosis/Prescription score Counseling Role Play: Communication/Interpersonal Skills score Training Role Play: Structure Training Role Play: Motivation Maintenance Situational Judgment Test: Total score

Figure 4.3. Model of second-tour NCO performance, based on LVII sample performance measures.

Analysis Steps

If the latent structure of performance is consistent across replications within organizational level and differs as expected (given the limitations of measurement) across levels (i.e., first tour vs. second tour), then it is reasonable to expect that the observed correlations of performance with performance, or performance factors with performance factors, would show the expected convergent and divergent relationships. That is, a particular performance factor measured at time one should have a higher correlation with itself at time two than it does with other performance factors at time two.

To examine these patterns of correlations for the Project A/Career Force Project Longitudinal Validation Sample, three basic intercorrelation matrices were computed.

(1) End-of-Training Performance vs. First-Tour Performance (10 x 9 matrix)

EOT measures:

6 EOT performance factor scores
- Effort and Technical Skill (ETS)
- Maintaining Personal Discipline (MPD)
- Physical Fitness and Military Bearing (PFB)
- Leadership Potential (LEAD)
- Basic Knowledge score (Base)
- Technical Knowledge score (Tech)

1 The sum of the Effort and Technical Skill and the Leadership scores (EOT:ELS)
1 The sum of the "can do" factors (EOT:Tech + EOT:Base)
1 The sum of the "will do" scores (EOT:ETS + EOT:LEAD + EOT:MPD + EOT:PFB)
1 The grand sum of all six factor scores (EOT: Total)

LVI measures:

5 First-tour performance factor scores
- Core Technical Proficiency (CTP)
- General Soldiering Proficiency (GSP)
- Effort and Leadership (ELS)
- Maintaining Personal Discipline (MPD)
- Physical Fitness and Military Bearing (PFB)

1 The sum of the "can do" scores (LVI:CTP + LVI:GSP)
1 The sum of the "will do" scores (LVI:ELS + LVI:MPD + LVI:PFB)
1 The grand sum of all five factor scores (LVI:Total)
1 The single scale rating of NCO potential (LVI:NCO)

(2) First-Tour Performance vs. Second-Tour Performance (9 x 11 matrix)

LVI measures:

9 As listed above for Matrix (1)

LVII measures:

6 Second-tour performance factor scores
- Core Technical Proficiency (CTP)
- General Soldiering Proficiency (GSP)

- Achievement and Effort (AE)
- Maintaining Personal Discipline (MPD)
- Physical Fitness and Military Bearing (PFB)
- Leadership (LEAD)

1 The sum of the LVII Achievement and Effort score and the LVII Leadership score

1 The LVII Leadership score computed without the Situational Judgment Test (a paper-and-pencil test; all other Leadership components are from ratings).

1 The sum of the "can do" factors

1 The sum of the "will do" factors

1 The grand sum of all six factor scores

(3) End-of-Training Performance vs. Second-Tour Performance (10 x 11 matrix)

EOT measures: As listed for matrix (1)
LVII measures: As listed for matrix (2)

Each matrix was calculated by computing the intercorrelations within each MOS and then averaging over MOS. All correlations are corrected for restriction of range by using a multivariate correction that treated the six EOT performance factors as the "implicit" selection variables on the grounds that, in comparison to other incidental selection variables, these factors would have the most to do with whether an individual advanced in the organization.

Making the corrections in this way means that the referent population consists of all the soldiers in the LV sample who had completed their training course. This is the population for which we would like to estimate the prediction of performance with performance, and it is the population for which the comparison of the validities of the experimental predictor tests and training criteria as predictors of future performance is the most meaningful. So long as the implicit selection variables are the best available approximation to the explicit selection variables, the corrected coefficients will be a better estimate of the population values than the uncorrected coefficients, but they will still be an underestimate (Linn, 1968). Since the degree of range restriction from EOT to LVI is not very great, the effects of the corrections were not very large.

RESULTS

The correlations of training performance with first-tour performance are shown in Table 4.2; the correlations of first-tour performance with second-tour performance are shown in Table 4.3; and the correlations of training performance and second-tour performance are shown in Table 4.4.

In general, there are substantial correlations of performance with performance. Performance in training does predict performance as a first-tour job incumbent, and performance in the first tour of duty does predict performance in the second tour after reenlistment. Performance in training also predicts performance during the second tour, approximately 5-6 years

Table 4.2

Zero-Order Correlations of Training Performance (EOT) Variables With First-Tour Job Performance (LVI) Variables: Weighted Average Across MOS

	EOT Variables[a]									
LVI Variables	EOT:TECH	EOT:BASE	EOT:ETS	EOT:MPD	EOT:PFB	EOT:LEAD	EOT:ELS	EOT:CAN	EOT:WILL	EOT:TOT
Core Technical Proficiency (CTP)	**.482** 3857	.380 3582	.217 3843	.153 3843	.049 3843	.180 3843	.208 3843	.475 3582	.180 3843	.485 3535
General Soldiering Proficiency (GSP)	.493 3857	**.452** 3582	.230 3843	.171 3843	.043 3843	.162 3843	.203 3843	.526 3582	.181 3843	.534 3535
Effort and Leadership (ELS)	.209 3795	.167 3525	.354 3783	.250 3783	.277 3783	.353 3783	**.376** 3783	.208 3525	.365 3783	.251 3479
Maintain Personal Discipline (MPD)	.174 3908	.136 3633	.310 3894	**.355** 3894	.214 3894	.272 3894	.307 3894	.170 3633	.340 3894	.211 3586
Physical Fitness and Bearing (PFB)	-.011 3908	-.016 3633	.262 3894	.127 3894	**.444** 3894	.308 3894	.307 3894	-.015 3633	.330 3894	.031 3586
"Can Do" Performance Composite (CAN)	.530 3857	.451 3582	.245 3843	.177 3843	.050 3843	.187 3843	.226 3843	**.545** 3582	.197 3843	.555 3535
"Will Do" Performance Composite (WILL)	.167 3795	.128 3525	.386 3783	.302 3783	.373 3783	.389 3783	.413 3783	.163 3525	**.427** 3783	.216 3479
Total Performance Composite (TOT)	.388 3741	.322 3471	.407 3729	.314 3729	.298 3729	.379 3729	.416 3729	.394 3471	.413 3729	**.438** 3425
NCO Potential Rating <Supv> (NCO)	.165 3458	.140 3458	.309 3444	.224 3444	.259 3444	.306 3444	.327 3444	.169 3458	.324 3444	.208 3397

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation. Correlations between matching variables are in bold.
[a] See full list in matrix (1) under "Analysis Steps."

later. This is true both for the variables measured with standardized tests and job samples and for the variables assessed via peer ratings.

There is also a reasonable pattern of convergent and divergent validity across performance factors, even without correcting these coefficients for attenuation and thereby controlling for the effects of differential reliability. The greatest departure from the expected pattern is found in the differential correlations of the two "can do" test-based factors (i.e., CTP and GSP). Current CTP does not always correlate higher with future CTP than current GSP correlates with future CTP. The correlation patterns for the "will do" factors, which are based largely on ratings, virtually never violate the expected pattern, even when peer ratings during training are being correlated with supervisory ratings obtained during the second tour.

The one possible exception to the consistent results for the "will do" factors is the predictability of the leadership performance factor for second-tour personnel. This component of NCO performance is predicted by almost all components of past performance. In part, this is because the LVII Leadership factor includes the Situational Judgment Test as one of its component parts. That is, the correlations of LVII Leadership with LVI CTP and ELS are reduced when the SJT is removed from the LVI Leadership score. However, they do not

Table 4.3

Zero-Order Correlations of First-Tour Job Performance (LVI) Variables With Second-Tour Job Performance (LVII) Variables: Weighted Average Across MOS

	LVI Variables[a]								
LVII Variables	LVI:CTP	LVI:GSP	LVI:ELS	LVI:MPD	LVI:PFB	LVI:CAN	LVI:WILL	LVI:TOT	LVI:NCO
Core Technical Proficiency (CTP)	**.440** 412	.413 412	.249 400	.078 413	.015 412	.449 412	.181 400	.375 397	.230 379
General Soldiering Proficiency (GSP)	.511 412	**.569** 412	.219 400	.085 413	-.008 412	.578 412	.157 400	.440 397	.220 379
Achievement and Effort (AE)	.103 390	.167 390	**.450** 377	.280 390	.319 390	.150 390	.464 377	.470 374	.412 353
Leadership (LEAD)	.359 344	.411 344	**.379** 333	.272 343	.169 342	.421 344	.365 333	.517 332	.378 319
Leadership Minus SJT Score	.264 348	.310 348	**.372** 337	.249 347	.233 346	.321 348	.380 337	.467 336	.398 322
Achievement, Effort and Leadership	.275 333	.335 333	**.471** 322	.292 332	.264 331	.336 333	.459 322	.620 321	.444 307
Maintain Personal Discipline (MPD)	-.044 406	.038 406	.116 393	**.257** 406	.166 406	-.002 406	.211 393	.164 390	.114 370
Physical Fitness and Bearing (PFB)	-.026 392	-.013 392	.220 379	.135 392	**.460** 392	-.022 392	.333 379	.250 376	.265 356
"Can Do" Performance Composite (CAN)	.520 412	.533 412	.259 400	.097 413	.010 412	**.562** 412	.193 400	.452 397	.260 379
"Will Do" Performance Composite (WILL)	.141 321	.190 321	.370 310	.295 320	.347 319	.182 321	**.433** 310	.445 309	.404 296
Total Performance Composite (TOT)	.336 313	.357 313	.381 302	.240 312	.252 311	.375 313	.394 302	**.521** 301	.423 289

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation. Correlations between matching variables are in bold.
[a] See full list in matrix (1) under "Analysis Steps."

reduce to zero, and there is still a general tendency for LVII Leadership to be predicted by almost all aspects of LVI performance.

For reference purposes the intercorrelations of each set of criterion factors, averaged over MOS, are shown in Tables 4.5, 4.6, and 4.7. As has been shown a number of times before, the intercorrelations for the "can do" factors and among the "will do" factors are relatively high and the cross-correlations between them are much lower. Given the high intercorrelations for the "will do" factors, the convergent and divergent relationship they exhibit is even more striking.

Table 4.4

Zero-Order Correlations of Training Performance (EOT) Variables With Second-Tour Job Performance (LVII) Variables: Weighted Average Across MOS

	EOT Variables[a]									
LVII Variables	EOT:TECH	EOT:BASE	EOT:ETS	EOT:MPD	EOT:PFB	EOT:LEAD	EOT:ELS	EOT:CAN	EOT:WILL	EOT:TOT
Core Technical	**.479**	.413	.215	.147	.080	.174	.204	.484	.183	.480
Proficiency (CTP)	1014	960	1056	1056	1056	1056	1056	960	1056	936
General Soldiering	.488	**.429**	.192	.107	.064	.112	.155	.496	.139	.489
Proficiency (GSP)	1014	960	1056	1056	1056	1056	1056	960	1056	936
Achievement and	.098	.151	**.248**	.172	.189	.238	.258	.141	.250	.165
Effort (AE)	946	896	983	983	983	983	983	896	983	874
Leadership (LEAD)	.322	.387	.294	.191	.152	**.254**	.289	.396	.264	.416
	900	854	931	931	931	931	931	854	931	832
Leadership Minus	.202	.284	.293	.203	.187	**.276**	.302	.274	.284	.302
SJT Score	905	905	936	936	936	936	936	905	936	881
Achievement,	.238	.310	.281	.197	.168	.258	**.285**	.307	.268	.328
Effort and Leadership	856	811	886	886	886	886	886	811	886	791
Maintain Personal	.080	.086	.210	**.260**	.162	.210	.224	.091	.249	.117
Discipline (MPD)	1006	953	1045	1045	1045	1045	1045	953	1045	929
Physical Fitness	-.047	-.007	.123	.067	**.320**	.208	.183	-.032	.208	-.005
and Bearing (PFB)	967	916	1003	1003	1003	1003	1003	916	1003	893
"Can Do" Performance	.527	.457	.228	.145	.082	.160	.201	**.534**	.181	.530
Composite (CAN)	1014	960	1056	1056	1056	1056	1056	960	1056	936
"Will Do" Performance	.168	.221	.270	.218	.235	.281	.295	.215	**.297**	.242
Composite (WILL)	823	780	852	852	852	852	852	780	852	761
Total Performance	.375	.386	.301	.225	.208	.273	.303	.417	.297	**.434**
Composite (TOT)	805	762	831	831	831	831	831	762	831	743

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation. Correlations between matching variables are in bold.
[a] See full list in matrix (1) under "Analysis Steps."

To address the issue of whether information about past performance contributes unique variance to the prediction of future performance over that contained in measures of ability, personality, and interests, two types of hierarchical regressions were carried out. The first sequence specified that the order of entry would be past performance first, followed by the four factor scores from the ASVAB, and then followed by the eight composite scores from the AVOICE, and then the seven factor scores from the ABLE, in that order. The second sequence was similar to the first except that the order of past performance and ASVAB was reversed.

Table 4.5

Intercorrelations of Training Performance (EOT) Variables in the LVI Sample: Weighted Average Across MOS

	EOT:TECH	EOT:BASEC	EOT:ETS	EOT:MPD	EOT:PFB	EOT:LEAD
Technical Total Score	1.000 3911					
Basic Total Score	.425 3911	1.000 3911				
Effort and TEchnical Skill (ETS)	.185 3844	.174 3844	1.000 3897			
Maintain Personal Discipline (MPD)	.143 3844	.121 3844	.678 3897	1.000 3897		
Physical Fitness and Bearing (PFB)	-.011 3844	.039 3844	.626 3897	.436 3897	1.000 3897	
Leadership Potential (LEAD)	.142 3844	.112 3844	.760 3897	.605 3897	.657 3897	1.000 3897

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation.

Table 4.6

Intercorrelations of First-Tour Job Performance (LVI) Variables in the LVI Sample: Weighted Average Across MOS

	LVI:CTP	LVI:GSP	LVI:ELS	LVI:MPD	LVI:PFB
LVI: Core Technical Proficiency (CTP)	1.000 3910				
LVI: General Soldiering Proficiency (GSP)	.590 3910	1.000 3910			
LVI: Effort and Leadership (ELS)	.294 3793	.282 3793	1.000 3847		
LVI: Maintain Personal Discipline (MPD)	.212 3907	.218 3907	.615 3847	1.000 3961	
LVI: Physical Fitness and Bearing (PFB)	.069 3907	.050 3907	.503 3847	.370 3958	1.000 3897

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation.

Table 4.7

Intercorrelations of Second-Tour Job Performance (LVII) Variables in the LVII Sample: Weighted Average Across MOS

	LVII:CTP	LVII:GSP	LVII:AE	LVII:LDR	LVII:MPD	LVI:PFB
LVII: Core Technical Proficiency	1.000 1458					
LVII: General Soldiering Proficiency	.544 1458	1.000 1458				
LVII: Achievement and Effort (AE)	.299 1312	.276 1312	1.000 1366			
LVII: Leadership (LDR)	.467 1237	.496 1237	.544 1208	1.000 1263		
LVII: Maintain Personal Discipline (MPD)	.139 1386	.152 1386	.536 1366	.411 1263	1.000 1443	
LVII: Physical Fitness and Bearing (PFB)	.094 1339	.091 1339	.432 1318	.289 1223	.359 1392	1.000 1392

Note. Corrected for range restriction. Pairwise *N*s are printed below each correlation.

The results of such a hierarchical analyses for the prediction of first-tour performance from training performance and the prediction of second-tour performance from first-tour performance are shown in Tables 4.8 through 4.11.

The general findings illustrated in these tables seem relatively clear. Both past performance and measured abilities add unique variance to the prediction of future performance. Knowledge of past performance adds relatively more, in comparison to trait measures, to the prediction of the "will do" components of future performance than to the task performance, or "can do" components. Conversely, the cognitive ability measure (ASVAB) adds more to the prediction of future performance on the "can do" factors while the ABLE adds relatively more to the prediction of the "will do" components.

SUMMARY CONCLUSIONS

Even though the intercorrelations of the performance factors identified at each of these organizational levels are reasonably high, the results of the above analyses suggest considerable construct validity for the measures of the factors. Past performance does predict future performance, and it does so with a considerable degree of convergent, divergent validity across the major components of performance.

Table 4.8

Multiple Correlations for Predicting First-Tour Job Performance (LVI) Criterion Factors From Training Criteria and Predictor Measures When the Training Performance (EOT) Variable is Matched With its Job Performance Counterpart, and Incremental Validities Obtained Where ASVAB Factors (A4), AVOICE, and ABLE are Added: Weighted Average Across MOS

First-Tour Job Performance Factor (LVI)	EOT [1][a]	EOT + A4 [5][a]	EOT + A4 + AVOICE [13][a]	EOT + A4 + AVOICE + ABLE [20][a]
Core Technical Proficiency (CTP)	.489	.538 (.536)	.539 (.534)	.541 (.533)
General Soldiering Proficiency (GSP)	.451	.577 (.576)	.579 (.575)	.584 (.577)
Effort and Leadership (ELS)	.379 (.378)	.416 (.412)	.425 (.416)	.431 (.419)
Maintaining Personal Discipline (MPD)	.356	.370 (.366)	.378 (.369)	.400 (.386)
Physical Fitness and Bearing (PFB)	.444	.455 (.452)	.465 (.458)	.492 (.483)

Note. Corrected for range restriction. Numbers in brackets are the number of predictor scores entering prediction equations. Numbers in parentheses are multiple Rs adjusted for shrinkage (Rozeboom Formula 8).

N = 3,523.

[a] Each job performance factor was matched with its corresponding training performance factor except for LVI:ELS, which was matched with the weighted sum of EOT:Effort and Technical Skill and EOT:Leadership Potential. Therefore, the number of predictors entering the equations for LVI:ELS is the number in brackets + 1.

Table 4.9

Multiple Correlations for Predicting First-Tour Job Performance (LVI) Criterion Factors From ASVAB (A4) and Incremental Validities Obtained When the Appropriate EOT Performance Factor, AVOICE, and ABLE are Added: Weighted Average Across MOS

First-Tour Job Performance Factor (LVI)	A4 [4]	A4 + EOT [5][a]	A4 + AVOICE [12]	A4 + AVOICE + EOT [13][a]	A4 + AVOICE + ABLE [19]	A4 + AVOICE + ABLE + EOT [20][a]
Core Technical Proficiency (CTP)	.467 (.464)	.538 (.536)	.471 (.464)	.539 (.534)	.477 (.467)	.541 (.533)
General Soldiering Proficiency (GSP)	.542 (.540)	.577 (.576)	.545 (.540)	.579 (.575)	.552 (.545)	.584 (.577)
Effort and Leadership (ELS)	.248 (.243)	.416 (.412)	.269 (.255)	.425 (.416)	.298 (.279)	.431 (.419)
Maintaining Personal Discipline (MPD)	.148 (.139)	.370 (.366)	.165 (.141)	.378 (.369)	.264 (.242)	.400 (.386)
Physical Fitness and Bearing (PFB)	.149 (.140)	.455 (.452)	.203 (.184)	.465 (.458)	.339 (.323)	.492 (.483)

Note. Corrected for range restriction. Numbers in brackets are the number of predictor scores entering prediction equations. Numbers in parentheses are multiple Rs adjusted for shrinkage (Rozeboom Formula 8).

N = 3,523.

[a] Each job performance factor was matched with its corresponding training performance factor except for LVI:ELS, which was matched with the weighted sum of EOT:Effort & Technical Skill and EOT:Leadership Potential. Therefore, the number of predictors entering the equations for LVI:ELS is the number in brackets + 1.

Table 4.10

Multiple Correlations for Training Performance (EOT) Variables Matched With Second-Tour Job Performance (LVII) Variables, and Incremental Validities When ASVAB Factors (A4), AVOICE, and ABLE are Added: Weighted Average Across MOS

Second-Tour Job Performance Factor (LVII)	EOT [1][a]	EOT + A4 [5][a]	EOT + A4 + AVOICE [13][a]	EOT + A4 + AVOICE + ABLE [20][a]
Core Technical Proficiency (CTP)	.482	.557 (.553)	.569 (.559)	.571 (.555)
General Soldiering Proficiency (GSP)	.418	.596 (.593)	.606 (.597)	.610 (.596)
Effort and Achievement + Leadership (ELS)	.387 (.384)	.434 (.426)	.446 (.429)	.471 (.448)
Achievement and Effort (AE)	.250	.292 (.281)	.305 (.279)	.330 (.291)
Leadership (LDR)	.257	.498 (.493)	.514 (.501)	.536 (.518)
Maintaining Personal Discipline (MPD)	.261	.272 (.260)	.300 (.272)	.324 (.285)
Physical Fitness and Bearing (PFB)	.321	.333 (.324)	.371 (.350)	.407 (.378)

Note. Corrected for range restriction. Numbers in brackets are the number of predictor scores entering prediction equations. Numbers in parentheses are multiple Rs adjusted for shrinkage (Rozeboom Formula 8).

N = 1,525.

[a] Each job performance factor was matched with its corresponding training performance factor except for LVII:ELS, which was matched with the weighted sum of EOT:Effort and Technical Skill and EOT:Leadership Potential. Therefore, the number of predictors entering the equations for LVII:ELS is the number in brackets + 1.

Table 4.11

Multiple Correlations for Predicting Second-Tour Job Performance (LVII) Criterion Factors From ASVAB (A4) and Incremental Validities Obtained When the Appropriate EOT Performance Factor, AVOICE, and ABLE are Added: Weighted Average Across MOS

Second-Tour Job Performance Factor (LVII)	A4 [4]	A4 + EOT [5][a]	A4 + AVOICE [12]	A4 + AVOICE + EOT [13][a]	A4 + AVOICE + ABLE [19]	A4 + AVOICE + ABLE + EOT [20][a]
Core Technical Proficiency (CTP)	.522 (.518)	.557 (.553)	.537 (.526)	.569 (.559)	.539 (.522)	.571 (.555)
General Soldiering Proficiency (GSP)	.588 (.585)	.596 (.593)	.599 (.590)	.606 (.597)	.603 (.589)	.610 (.596)
Effort and Achievement + Leadership (ELS)	.366 (.360)	.434 (.426)	.384 (.366)	.446 (.429)	.425 (.400)	.471 (.448)
Achievement and Effort (AE)	.193 (.179)	.292 (.281)	.217 (.179)	.305 (.279)	.258 (.208)	.330 (.291)
Leadership (LDR)	.456 (.452)	.498 (.493)	.477 (.464)	.514 (.501)	.510 (.492)	.536 (.518)
Maintaining Personal Discipline (MPD)	.111 (.084)	.272 (.260)	.176 (.125)	.300 (.272)	.242 (.187)	.324 (.285)
Physical Fitness and Bearing (PFB)	.128 (.106)	.333 (.324)	.233 (.198)	.371 (.350)	.333 (.297)	.407 (.378)

Note. Corrected for range restriction. Numbers in brackets are the number of predictor scores entering prediction equations. Numbers in parentheses are multiple Rs adjusted for shrinkage (Roseboom Formula 8).

N = 1,525.

[a] Each job performance factor was matched with its corresponding training performance factor except for LVII:ELS, which was matched with the weighted sum of EOT:Effort and Technical Skill and EOT:Leadership Potential. Therefore, the number of predictors entering the equations for LVII:ELS is the number in brackets + 1.

Chapter 5
PREDICTION OF FIRST-TERM MILITARY ATTRITION USING PRE-ENLISTMENT PREDICTORS

Rodney A. McCloy and Ani S. DiFazio

INTRODUCTION

Soldiers who fail to complete their contracted first term of service are costly to the military. The cost of first-term attrition was conservatively estimated to be $200 million (in 1989 dollars) by Klein, Hawes-Dawson, and Martin (1991). Direct costs include lost investments such as training and recruiting costs, and compensation costs in the form of salary during enlistment and subsequent unemployment costs (e.g., Laurence, 1993; McCloy et al., 1992). Laurence further noted that attrition also leads to significant indirect costs, both to the military (in the form of lowered morale and force instability) and to the individual (in the form of possibly reduced future employment opportunities and earning potential). As a result, significant benefits can accrue from better understanding the precursors of attrition and using this information to select those recruits who are less likely to exit the military prematurely.

The Services currently use high school diploma graduate status as an indicator of a recruit's chances of completing his or her first term, a practice spurred by an initial Air Force technical report (Flyer, 1959) and since justified by years of supporting evidence (e.g., Department of Defense, 1985). As Laurence (1993) stated, the Flyer study "was only the first in a very long line of research to conclude that high school graduates have lower attrition rates [than non-diploma graduates] . . . [and] similar findings have been echoed in countless reports over the past 30 plus years" (p. 5).

In contrast, very little relationship has been found between measures of cognitive ability and attrition. High school graduates scoring in the lowest part of the distribution on the Armed Forces Qualification Test (AFQT) have lower rates of attrition than non-graduates in the uppermost part of the distribution. Laurence (1993) suggested that the relationship between first-term attrition and diploma status might be due to differences between the two groups on various non-cognitive characteristics. These characteristics, in turn, are typically assessed with instruments such as temperament, interest, or biodata inventories.

Indeed, although high school diploma status is the best single predictor of first-term attrition, biodata instruments have demonstrated incremental validity (e.g., Steinhaus, 1988; Trent, 1993; White, Nord, Mael, & Young, 1993). Further, research from the Project A/Career Force program (e.g., Campbell & Zook, 1991, 1994a) has demonstrated the validity of non-cognitive measures for predicting the volitional, or "will do," dimensions of Army job performance (e.g., effort, physical fitness and military bearing, personal discipline) -- dimensions that may impact attrition.

Objectives

The present research has two primary goals. The first is to determine the relationship between first-term attrition and the three non-cognitive

predictor measures: the Assessment of Background and Life Experiences (ABLE), the Army Vocational Interest Career Examination (AVOICE), and the Job Orientation Blank (JOB). A critical research question concerns the incremental contributions of these measures to the prediction of attrition over the information provided by the Armed Services Vocational Aptitude Battery (ASVAB) and high school diploma graduate (HSDG) status.

The second goal is to use all the available predictor data, as described above, to develop a specific predictor composite for attrition that could be used to select applicants who have higher probabilities for completing their first term of service.

The relationship between first-term attrition and the pre-enlistment predictors was addressed using proportional hazards modeling (Cox, 1972), a form of event history analysis (cf. Allison, 1984). Event history analysis allows a researcher to model whether an event occurs, and if so, when it occurs. In the present research, the event is attrition from the Army during the first term of enlistment. Specifically, a proportional hazards model allows one to determine the relationship between one or more predictor variables and the rate at which events occur over time.

Analytic Difficulties With Event Data

Researchers who describe events as dependent variables must be aware that event data bring with them certain analytic difficulties (Allison, 1984; Singer & Willett, 1991). Three problems frequently appear. First, the presence of time-varying independent variables often complicates the analysis process (Allison, 1984, pp. 10-11). Time-varying independent variables do not appear in every event study, but it is often sensible to consider them, given that the observation of events naturally occurs over time, and a person's status on many potential predictors may be expected to change in a meaningful way during the elapsed time.

Second, in the present research some soldiers leave for reasons other than those defined as first-term attrition (e.g., soldiers going to Officer Candidate School). These soldiers did not experience the event during the time they were observed, but they were not observed for the entire first term. Such observations are said to be "censored."

There are actually several types of censoring (cf. Lawless, 1982). The censoring in the above example is typically termed "right censoring." As time moves left to right on a time line, a person's observation period runs along the time line until it hits a border on the right side: the point where the observation period ends. Any events or non-events after (i.e., to the right of) this demarcation are unknown to the researcher (i.e., they are censored).

Two types of right-censored observations occur in the current data. First, those soldiers who complete their term of enlistment are right-censored at $t = 3$ (or 4) years. One cannot be sure that these soldiers did not leave the military two years, two months, or even two days after the end of the contracted time, which may or may not be the end of their first term (given the possibility of contract extensions). All that can be said is that they did not exit prematurely during the first term. Second, soldiers who exited the Army for reasons not defined as attrition (e.g., Officer Candidate School) are also right-censored, but at the point they exited the enlisted ranks.

Similar to those soldiers who did not attrit, all that can be said of these soldiers is that they had not yet exited prematurely as of time $t < 3$ (or 4) years (the end of the contracted enlistment). The question is how to handle the censored observations. Failure to treat censored observations appropriately can lead to biased results.

Third, it is clear that the distribution of the events over time in no way approximates a normal distribution. This is the rule rather than the exception for event data. Hence, familiar analytic techniques based upon normal theory are inappropriate for the analysis of event data.

Several strategies aimed at accommodating these characteristics of event data have been tried (cf. Singer & Willett, 1991). For example, one might perform an ordinary least-squares (OLS) regression on the dichotomous dependent variable ATTRITION, indicating whether the event did or did not occur as of a specified point in time (e.g., $t = 2$ years). The dichotomous variable ignores the timing of the events, focusing instead on whether or not an event occurred during the observation period. Using OLS procedures with a dichotomous variable of this sort creates a number of undesirable properties such as the possibility of out-of-range predicted values for the dependent variable and violation of assumptions regarding the variance of the dependent variable (e.g., Collett, 1991).

A second option would be to ignore the timing of the events by performing a logistic regression analysis of the same dichotomous dependent variable. This type of analysis avoids many of the pitfalls of the OLS approach (e.g., the predicted values of the criterion will not lie outside acceptable bounds) and is reasonable if (a) few individuals exit the study and (b) the timing of event occurrence holds little interest for the researcher. Nevertheless, the approach does not easily accommodate time-varying independent variables or censored observations. Individuals who exited the study typically will be deleted from this design because they do not qualify as an event, and labeling them as a non-event is inaccurate. Clearly, the greater the number of exiting observations, the greater the potential for bias in the parameter estimates of the regression model if they are excluded from the analysis. (In the present data, the number of mid-term censored observations is quite small, suggesting that the potential for bias in a logistic regression analysis is likely to be small.)

If the occurrence of events over time is of interest, a second strategy is to model the time to the event by changing the dependent variable into a continuous variable, such as "days of service." This approach, however, also has drawbacks. Consider the issue of censoring, in particular. The exiting observations again must be excluded from the analysis, and there is a ceiling effect on the dependent variable for the observations that carry through the entire observation period without experiencing an event that is directly influenced by the often arbitrary duration of the study period. Further, the ceiling is inaccurate because it represents the time elapsed before the study ended rather than an event time.

A class of methods is available that accommodates all of the issues just mentioned. These methods have been used for many years in engineering (e.g., Lawless, 1982) and biostatistics (e.g., Kalbfleisch & Prentice, 1980) but have gained attention only recently in the psychological literature (e.g., Fichman, 1988, 1989; Harrison & Hulin, 1989; Morita, Lee, & Mowday, 1989; Singer &

Willett, 1991). The methods carry several labels, including "failure-time analysis" from the engineering literature and "survival analysis" from the biostatistical literature. Perhaps the most general term for these methods is "event history analysis" (Allison, 1984).

Event history analysis allows a researcher to model whether an event occurs, and if so, when it occurs. In many ways, event history models have much in common with traditional analytic strategies. For example, event history analyses generate both descriptive and inferential statistical information. Group differences in event occurrence can be tested, and statistical models relating independent variables to event occurrence can be developed. Nevertheless, the mathematics of event history analyses is more complex than the mathematics of correlational analytic methods. The basic elements of event history models are described below. Formal treatments of the mathematics of event history models appear in Kalbfleisch and Prentice (1980) and Lawless (1982).

Functions: The Building Blocks of Event History Models

Event history models involve functions, and the function values are evaluated across time. The two primary functions used in event history analyses are the survivor function and the hazard function.

Survivor Function

The survivor function, S(t), describes the probability that an individual will survive at least until time t without experiencing the event in question. S(t) is a monotonic, non-increasing (typically decreasing) function. In this respect, it is essentially a reverse cumulative distribution function, cumulating across time the proportion of observations that have yet to experience the event.

Figure 5.1 contains a plot of survival curves for the first-term attrition of high school graduates and non-graduates who enlisted for three years into the Army Military Occupational Specialty (MOS) 11B (Infantryman). The curves indicate that the survival rate for graduates is much higher than for non-graduates. For example, it takes just under two years to reach the point where 80 percent of the original sample of graduates remain, whereas it takes just 10 months to winnow the sample of non-graduates to 80 percent of its original size.

Hazard Function

Because S(t) is a monotonically non-increasing function, its shape remains relatively unchanged, regardless of the rate at which events occur over time. Thus, the survivor function might appear relatively uninformative, because not all events share the same pattern of occurrence. For example, the probability of dying increases with a function of time from about age 30 onward (Kalbfleisch & Prentice, 1980), whereas the probability of leaving the military increases rapidly during the first three months of service and decreases to a relatively stable rate thereafter. The function describing the distribution of event occurrence across time is h(t), the hazard function. The definition of the hazard function depends upon whether time is measured in discrete units or continuously.

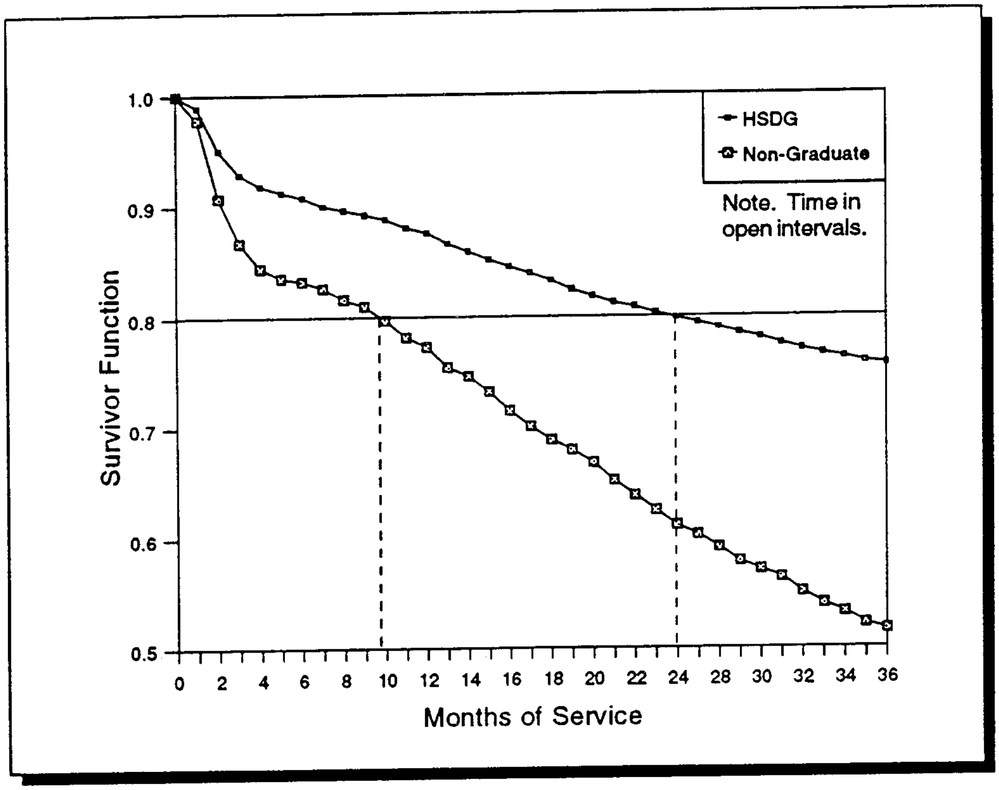


Figure 5.1. Survivor curves for high school diploma graduates and non-graduates in MOS 11B with 3-year terms of enlistment.

Discrete time. For discrete time, h(t) represents the probability an individual will experience an event during a specified time interval, given that the individual is at risk for experiencing the event. Hence, the hazard is a conditional density function. Calculation of the discrete-time hazard depends upon two quantities: (a) the number of individuals who experience the event during the interval, divided by (b) the number of individuals who are at risk for experiencing the event during the interval -- what Allison (1984) labeled the risk set. For single events, the risk set steadily decreases as individuals either experience the event or are censored.

The risk set and its use in calculating the discrete time hazard demonstrate how an event history model makes optimal use of data from censored observations. Consider a researcher investigating what variables predict the event of divorce within the first five years of marriage. Assume that 500 couples appear in the original sample at time t = 0 and that observations of marital status are made twice a year. At the first observation period of six months, 5 of the 500 couples will have divorced. Hence, for the first six-month period, the discrete time hazard is 5/500 = .01. During the second time interval, 25 couples obtain divorces and 10 couples exit the study while married (i.e., they are censored observations). The risk set is now 495 rather than 500, because the five couples who divorced during the first six months are no longer part of the sample. Hence, h(t) for the second time interval is 25/495 = .05. Note that the censored observations contribute to the risk set for interval two but are not considered events, because they did

not divorce. For the third time interval, however, the risk set will be (495 - 35) = 460. Thus, the censored observations do not contribute to the risk set for the third interval. This example demonstrates how the data for censored observations are used correctly and optimally in event history models, contributing to the calculation of the hazard rate (via the risk set) for the amount of time the observations are in the study.

Discrete-time event history models provide an easily understood interpretation of the hazard function. In addition, discrete-time models can be estimated with standard logistic regression software programs. For discussions and applications of discrete-time models, see Singer and Willett (1993) and Willett and Singer (1993, in press).

Continuous time. For continuous time, h(t) is defined as a mathematical limit that describes the instantaneous rate at which events occur. Formally,

$$h(t) = \lim_{\Delta t \to 0} \frac{p(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \qquad \textbf{(1)}$$

where T is the time of the event (Kalbfleisch & Prentice, 1980, p. 6). The numerator of the limit is the discrete time hazard rate when $\Delta t = 1$. For the continuous time hazard rate, the probability is divided by Δt, the length of the interval; the limit is evaluated as this interval becomes smaller. The continuous time hazard is not a probability, because it can take on values greater than one, having a range of $[0,+\infty)$ (Allison, 1984).[1] The natural logarithm of the hazard rate typically serves as the dependent variable in event history models, as will be shown in a later section. One way of conceptualizing the continuous-time hazard rate for the current situation is as the relative rate at which people are leaving the Army at any given point in time.

Unlike the survivor function, h(t) can take any shape. Figure 5.2 shows a plot of the hazard functions corresponding to Figure 5.1. As mentioned above, these plots indicate that the rate of first-term attrition is greatest during the first six months of service, after which it decreases to a relatively steady rate.

---

[1] The survivor and hazard functions are related. Formally, the hazard rate is defined as

$$h(t) = \frac{f(t)}{S(t)}$$

where f(t) is the probability density function of the event times and S(t) is the survivor function. Note that

$$f(t) = -\frac{dS(t)}{dt}$$

because S(t) is a reverse cumulative distribution function (Kalbfleisch & Prentice, 1980).

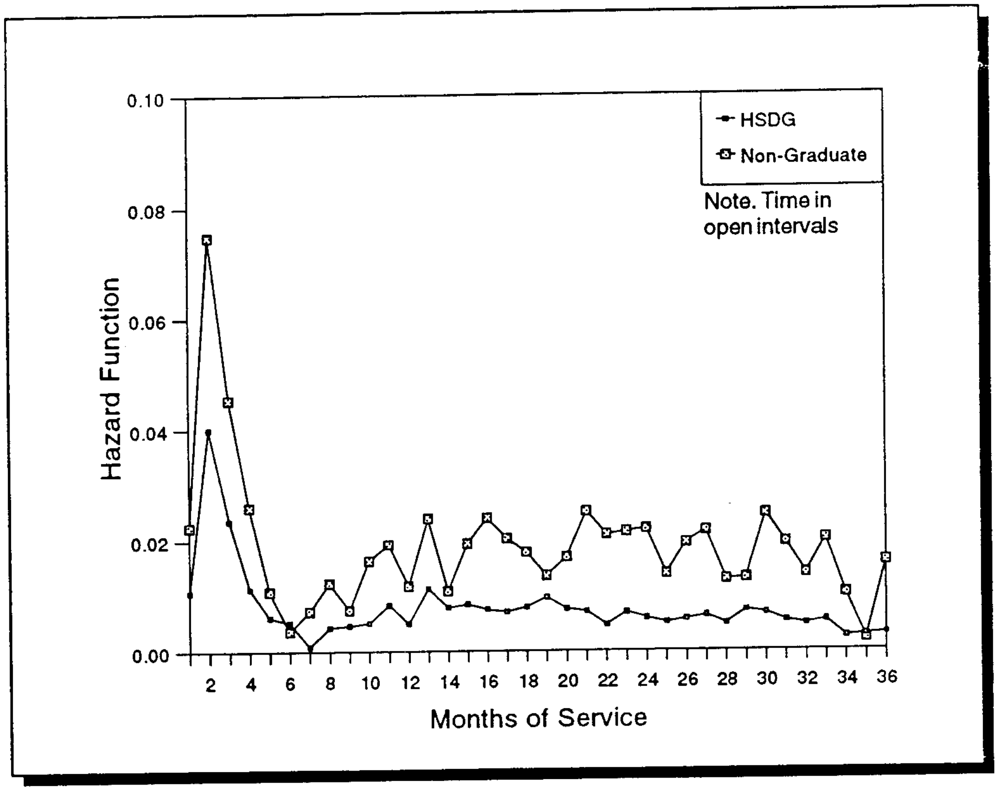


Figure 5.2. Hazard curves for high school diploma graduates and non-graduates in MOS 11B with 3-year terms of enlistment.

Descriptive Statistical Information

When conducting correlational analyses, one typically wishes to determine the characteristics of (i.e., describe) the sample. The most common means of doing this is to calculate the means, standard deviations, ranges, and so on of the measures administered to that sample. In event history analysis, descriptive information is provided by plots of the survivor and hazard functions, which may be obtained for the total sample or for subgroups. Corresponding information on the number of events that occurred for the subgroup and the number of censored observations is also useful.

Another common step in correlational analyses is to test for mean differences on one or more measures between various groups in the analysis sample. A similar strategy applies to event history analyses, but rather than testing subgroup differences on specific measures, the subgroup survivor functions are compared to see if the event behavior differs by subgroup. For example, the survivor functions in Figure 5.1 could be compared to determine if the groups' survival behavior differs (and it does at $p < .0001$). Subgroup differences may also be investigated while controlling for the effects of one or more variables the investigator thinks might be related to survival times.

The differences in the survivor functions are tested using versions of log-rank tests, the most commonly reported being the Mantel-Cox (or Savage) statistic, the generalized Wilcoxon (or Breslow) statistic, and the Tarone-Ware statistic. These log-rank statistics essentially test the difference

between the observed number of events in each of the groups and the number of events that would be expected if the survivor functions of all the groups were equivalent. The various statistics differ primarily in the manner in which they weight observations. For example, all observations are weighted equally in the calculation of the Mantel-Cox statistic, whereas earlier observations receive greater weight than later observations for the generalized Wilcoxon statistic, thus giving less weight to later events when fewer individuals remain in the risk set.

Log-rank tests are asymptotically distributed as χ^2 with (g-1) degrees of freedom, where g is the number of groups. Hence, their use in small samples should be approached with caution.

Inferential Statistical Information

After examining the descriptive information, one may wish to develop predictive models of the dependent variable(s) of interest. For normally distributed continuous variables, multiple regression analysis can be used to generate prediction equations relating one or more predictors to the criterion. These equations can be used to obtain predicted scores on the criterion for each observation having predictor data.

A similar modeling strategy can be used in event history analysis, but the criterion variable, method of estimation, fit statistics, and predicted values take different forms. Numerous models are available. Allison (1984) distinguishes between parametric models, which impose a specific distribution on the event times; non-parametric models, which make few or no assumptions about the distribution of event times; and semi-parametric models, which assume a particular functional form for the regression parameters but do not impose any distribution on the event times. Detailed information on various parametric models is available in Kalbfleisch and Prentice (1980) and Lawless (1982). In this section, the semi-parametric proportional hazards model (Cox, 1972) will be discussed.

The Proportional Hazards Model

Since its introduction by Cox in 1972, the proportional hazards model has become one of the most widely used event history models. Formally, the model is:

$$\ln[h(t)] = \ln[h_0(t)] + \beta X \tag{2}$$

where ln is the natural logarithm, h(t) is the hazard rate, $h_0(t)$ is a baseline hazard rate that can take any form (i.e., it is non-parametric), β is a vector of regression coefficients, and **X** is a vector of predictor variables. Thus, the proportional hazards regression model relates a set of predictors to the (log) hazard of the event. Recall that h(t) has a range of $[0,+\infty)$. Taking the natural logarithm of the hazard provides an unbounded criterion variable with a range of $(-\infty, +\infty)$. The regression coefficients are estimated using a method devised by Cox called partial likelihood. The method has much in common with maximum likelihood estimation procedures.

The model gets it name from the fact that for any two individuals with covariate vectors X_1 and X_2, the ratio of their hazards is a constant value, k (i.e., the hazards are proportional). That is,

$$\frac{h_1(t)}{h_2(t)} = k \quad . \tag{3}$$

The proportionality assumption can be tested for the hazard functions of two groups by examining the significance of the parameter for a group dummy variable (δ) interacting with time. The following models are calculated:

$$\begin{aligned} \ln[h(t)] &= \ln[h_0(t)] + \beta_1\delta \\ \ln[h(t)] &= \ln[h_0(t)] + \beta_1\delta + \beta_2 X \end{aligned} \tag{4}$$

where h(t) is the hazard rate, $h_0(t)$ is the baseline hazard for the group(s) not modeled by δ, and X is the interaction between δ and the log of the event times. If β_2 is significant, then the ratio of the hazard plots is not constant, indicating that the hazard functions of the group modeled by δ and the other group(s) are not proportional.

There is no general consensus about the importance of testing the proportionality assumption. For example, Singer and Willett (1991) suggested that the proportional hazards assumption is virtually never justified and its plausibility should always be tested. In contrast, Allison (1984) stressed that the concept of proportional hazards is not an essential component of the model, given that (a) the hazards fail to be proportional as time-varying independent variables enter the model, and (b) even when the assumption is violated, the model often provides reasonable results. He argued that researchers would better spend their time reducing the likelihood of specification error in their models than worrying about possible violations of the proportionality assumption.

If the hazards prove not to be proportional, a stratified analysis may be conducted where "group" is the stratifying variable. The baseline hazards are allowed to differ for each group, but the regression parameters are assumed to be the same across groups. That is,

$$\begin{aligned} \textit{Group } 1\text{:} \quad \ln[h(t)] &= \ln[h_{01}(t)] + \beta_1 x_1 + \cdots + \beta_k x_k \\ \textit{Group } 2\text{:} \quad \ln[h(t)] &= \ln[h_{02}(t)] + \beta_1 x_1 + \cdots + \beta_k x_k \end{aligned} \tag{5}$$

where $h_{01}(t)$ and $h_{02}(t)$ are group-specific hazards that are not proportional to each other. This is somewhat analogous to estimating regression equations that allow group intercepts to differ while constraining the slopes to be equal across groups.

Whether the hazards are proportional for all groups or only within stratified groups, the effect of the independent variables on the hazard is to shift the baseline hazard up or down, depending on the values of the

variables. More formally, the effects of the independent variables are multiplicative on the hazard rate. This may be seen by taking the antilog of equation 2:

$$h(t) = h_0(t)\, e^{\beta x} \quad . \tag{6}$$

In a stratified analysis, an individual's hazard is shifted up or down relative to the baseline hazard for their group, whence equation 6 becomes

$$h(t) = h_{0i}(t)\, e^{\beta x} \quad . \tag{7}$$

with i being one of g groups (i = 1, . . ., g). The <u>magnitude</u> the hazard is shifted is the same for all individuals having identical **X** vectors, regardless of group membership. The <u>shape</u> of the hazard that is shifted, however, varies across groups (i.e., strata).

<u>Assessing the fit of proportional hazards models</u>. Determining the fit of event history models to the data could prove frustrating for those accustomed to conventional multiple regression analyses, because no measure akin to the coefficient of determination (R^2) is available, although approximations have been proposed (e.g., Harrell, 1980). Nevertheless, a statistic similar to the F test for the hypothesis that all the parameter coefficients are zero is available--the likelihood ratio test. The likelihood function is computed for two models: (a) the model in question with a select number of independent variables, or covariates, and (b) a model in which all parameters are zero. The difference in the logs of these likelihood values, multiplied by -2, is distributed asymptotically as χ^2, with degrees of freedom equal to the number of parameters in the model containing predictors. To facilitate the use of the likelihood ratio test, most statistical software packages provide the value of -2 Log L in the printout. One may then subtract these values for two nested models.

The likelihood ratio test can also be used to test whether the addition of one or more predictors to a given model significantly increases the fit of the model to the data. One simply multiplies the difference in the log-likelihoods for the two models by -2. As with the null model, the difference is distributed asymptotically as χ^2, with degrees of freedom equal to the difference in the number of parameters between the two models.

Note that the likelihood ratio test can be applied only to models that are nested (i.e., they are subsets of one another). This restriction, however, applies equally to tests of incremental R^2 that are available with linear multiple regression models.

<u>Cross-validity of event history models</u>. Another difficulty stemming from the lack of a fit statistic such as R^2 is that cross-validation does not have an immediate counterpart in event history analysis. No predicted scores are available for individuals that may be compared to actual values. Rather, the models generate predicted functions over time for each individual, which implies that each individual's predicted "score" depends upon time. Likewise, no shrinkage formulae are available for estimating the performance of the

sample coefficients in another sample or in the population. Nevertheless, McCloy (1993) has proposed a split-sample approach similar to that used in double cross-validation for estimating the cross-validity of event history models.

AN APPLICATION OF EVENT HISTORY ANALYSIS TO FIRST-TERM ATTRITION

Given the properties of event history analysis as described above, the proportional hazards model was used to analyze attrition data in the Project A/Career Force Longitudinal Validation sample. The specific objective was to determine the extent to which information from the Project A/Career Force Experimental Battery could add to the prediction of attrition over the full course of the first tour of duty.

Method

Subjects

The subjects were the roughly 49,000 first-term soldiers from the Project A/Career Force Longitudinal Validation (LV) sample (Campbell & Zook, 1994). These soldiers were administered a 4-hour Experimental Predictor Battery (Campbell & Zook, 1991) within 2 days of their entry into the Army. The data were collected over a 15-month period (20 Aug 86 through 30 Nov 87) at eight Reception Battalions. The jobs represented by the 21 MOS in the LV sample are listed in Table 5.1.

Table 5.1

Summary of the 21 Army MOS in the Project A/Career Force LV Sample

MOS	Name of Job	MOS	Name of Job
11B	Infantryman[a]	55B	Ammunition Specialist
12B	Combat Engineer[a]	63B	Light-Wheel Vehicle Mechanic
13B	Cannon Crewman[a]	67N	Utility Helicopter Repairer
16S	MANPADS Crewman[a]	71L	Administrative Specialist
19E[b]	M60 Armor Crewman[a]	76Y	Unit Supply Specialist
19K[b]	M1 Armor Crewman[a]	88M	Motor Transport Operator
27E	Tow/Dragon Repairer	91A	Medical Specialist
29E	Comm.-Electronics Radio Repairer	94B	Food Service Specialist
31C	Single Channel Radio Repairer	95B	Military Police
51B	Carpentry/Masonry Specialist	96B	Intelligence Analyst
54B	NBC Specialist		

[a] Combat group
[b] MOS 19E and 19K differ only in terms of equipment (i.e., type of tank).

Data checks for out-of-range values (e.g., impossible or incompatible entry and exit dates) reduced the sample to 48,308. Only those soldiers having complete data on the pre-enlistment measures were used to model attrition, yielding a final sample size of 31,032 soldiers.

Job groups. Rather than running analyses by MOS, analysts formed job groups by splitting the 21 MOS listed in Table 5.1 into two groups: Combat (MOS 11B, 12B, 13B, 16S, 19E, and 19K) and Non-Combat (all others). These two groups (C and NC) were subdivided by enlistment terms, as described below.

In many applications of event history analysis, the endpoints and length of the observation period are arbitrary, being driven by convenience or a desire to observe a certain number of events. First-term attrition, however, has clear starting and ending points, beginning when the soldier enters the military and ending at the conclusion of the enlistment term agreed upon on the enlistment contract. Army enlistment terms typically range from 2 to 6 years, with 3-year and 4-year terms being the most common. The present study contains only those soldiers who agreed to 3-year and 4-year terms.

Because the two terms of enlistment provide meaningful observation periods of different duration, the two MOS groups were split by term, yielding four analysis groups: soldiers in combat and non-combat MOS with 3-year and 4-year enlistment terms (C3, C4, NC3, and NC4, respectively). Note that all six Combat MOS appear in both enlistment term groups (C3 and C4) because they contain many soldiers with 3-year enlistment terms and many soldiers with 4-year terms. Similarly, three of the Non-Combat MOS (76Y, 94B, and 95B) appear in both Non-Combat enlistment term groups (NC3 and NC4). The MOS within each of the four job groups, the sample sizes for each, the number and percentage who experienced attrition, and the number and percentage who were censored observations are given in Tables 5.2 through 5.5.

Table 5.2

Sample Sizes, and Number and Percent of Attritions and Censored Observations for Soldiers With 3-Year Enlistments in Combat Jobs (C3)

		Events		Censored Observations	
MOS	N	N	Percent	N	Percent
11B	4,875	1,354	27.8	3,521	72.2
12B	1,131	295	26.1	836	73.9
13B	2,416	729	30.2	1,687	69.8
16S	410	140	34.1	270	65.9
19E	287	91	31.7	196	68.3
19K	570	128	22.5	442	77.5
Total	9,689	2,737	28.2	6,952	71.8

Table 5.3

Sample Sizes, and Number and Percent of Attritions and Censored Observations for Soldiers With 4-Year Enlistments in Combat Jobs (C4)

MOS	N	Events N	Events Percent	Censored Observations N	Censored Observations Percent
11B	4,196	1,397	33.3	2,799	66.7
12B	247	73	29.6	174	70.4
13B	1,227	405	33.0	822	67.0
16S	81	16	19.7	65	80.3
19E	123	47	38.2	76	61.8
19K	566	186	32.9	380	67.1
Total	6,440	2,124	33.0	4,316	67.0

Table 5.4

Sample Sizes, and Number and Percent of Attritions and Censored Observations for Soldiers With 3-Year Enlistments in Non-Combat Jobs (NC3)

MOS	N	Events N	Events Percent	Censored Observations N	Censored Observations Percent
54B	566	149	26.3	417	73.7
55B	198	73	36.9	125	63.1
71L	1,341	362	27.0	979	73.0
76Y	518	115	22.2	403	77.8
88M	769	247	32.1	522	67.9
91A	2,685	649	24.2	2,036	75.8
94B	1,579	547	34.6	1,032	65.4
95B	2,907	619	21.3	2,288	78.7
Total	10,563	2,761	26.1	6,952	73.9

Table 5.5

Sample Sizes, and Number and Percent of Attritions and Censored Observations for Soldiers With 4-Year Enlistments in Non-Combat Jobs (NC4)

		Events		Censored Observations	
MOS	N	N	Percent	N	Percent
27E	104	28	26.9	76	73.1
29E	188	50	26.6	138	73.4
31C	390	143	36.7	247	63.3
51B	319	88	27.6	231	72.4
63B	1,353	419	31.0	934	69.0
67N	277	55	19.9	222	80.1
76Y	874	275	31.5	599	68.5
94B	429	179	41.7	250	58.3
95B	406	99	24.4	307	75.6
Total	4,340	1,336	30.8	3,004	69.2

Measures

All measures were taken from the Project A/Career Force data base. The predictor measures were the same ones used in the LVI validation analyses.

Attrition. For this research, attrition is defined as a premature separation from first-term service for reasons that might be viewed negatively from the military perspective. Three critical components of this definition require clarification: (a) what is meant by premature, (b) which types of separation would be viewed negatively by the military, and (c) how to establish the time window of the first tour.

First, "premature" is defined as any length of service that is less than the tour to which recruits obligated themselves at enlistment. Second, as developed in another project sponsored by the Office of the Secretary of Defense, the Compensatory Screening Model (CSM)[2] was used to group separation types into "pejorative" and "non-pejorative" categories. Separations that the CSM identified as "pejorative," and that correspond to Knapp's (1993) Army separation behavior categories four and five were used to define negative separations. The separations that were used in this research are given in Table 5.6.

---

[2] The CSM is a selection procedure for identifying those non-high school diploma graduates who are more likely to complete at least two years of obligated service (cf. McBride, 1993).

Table 5.6

Behaviors Treated as Attrition in the Present Research[a]

- Failure to meet minimum behavioral or performance criteria (e.g., poor performance, disciplinary problems, adjustment problems)
- Medical conditions existing prior to enlistment or which indicate failure to conform to Service requirements (e.g., weight/body fat standards)
- Erroneous enlistment, underage
- Marriage, pregnancy, parenthood, family dependency/hardship
- Conscientious objector
- Desertion or imprisonment

[a] Based on Knapp (1993), Alternative Conceptualizations of Turnover.

Third, defining the first-term window of time is easy for soldiers who did not reenlist and for those who reenlisted once their initial term of obligation was completed: It is the time between accession and (first) separation from the Army. However, of the soldiers in the present dataset who reenlisted, 84.2 percent did so before, rather than after, their first term was completed. For this reason, identifying the length of the first tour for these soldiers is less clear.

The following rule was developed for these immediate reenlistments: (a) If the time between the basic active service date (the date establishing the beginning of a soldier's creditable active service; it is adjusted for prior service, AWOL, etc.) and separation from the reenlistment tour (or 30 Sep 92, the final observation date available in the data base) was at least as long as the initial enlistment term, then the first-term window of time was set to the enlistment term. Of the 10,342 soldiers who reenlisted during their first term, 10,248 (99.1%) had their time window set to their enlistment term -- that is, their total time in the military exceeded their initial obligation. (b) For the remaining 94 soldiers (0.9%) -- those whose time from accession to separation following reenlistment was less than the initial enlistment term (e.g., a soldier who initially enlisted for a three-year tour, reenlisted after six months, and then separated from the Army after only two years of service) -- the first-term window was set to the time between accession and the reenlistment date.

This treatment was chosen because we did not wish to include second-term (i.e., reenlistment) time as first-term time if the soldier exited prior to the contracted enlistment term. On the other hand, we wished to credit a soldier for completing his or her contracted enlistment, even if it was accomplished as part of an immediate reenlistment.

__High School Diploma Graduate Status (HSDG)__. A dummy variable was constructed indicating whether the individual was a high school diploma graduate (HSDG=1) or not (HSDG=0).

__Armed Services Vocational Aptitude Battery (ASVAB)__. The ten ASVAB subtests were combined into four composite scores: Quantitative, Speed, Technical, and Verbal (cf. Campbell, 1987; Waters, Barnes, Foley, Steinhaus, & Brown, 1988). These four composites were used in the present analyses.

__Assessment of Background and Life Experiences (ABLE)__. The ABLE, a 199-item temperament and biodata inventory, was designed to predict Army-relevant criteria, including attrition (Hough, Barge, & Kamp, 1987). Three response options are available for each item. The ABLE contains 11 substantive (i.e., content) scales and four response validity scales. The current analyses included the 11 substantive scales and the response validity scale measuring the Social Desirability of the soldier's responses. The 11 substantive scales are as follows:

Emotional Stability	Work Orientation
Self-Esteem	Internal Control
Cooperativeness	Energy Level
Conscientiousness	Dominance
Nondelinquency	Physical Condition
Traditional Values	

__Army Vocational Interest Career Examination (AVOICE)__. The AVOICE is an interest inventory that is based on the Air Force's Vocational Interest Career Examination (Peterson et al., 1990). It contains four sections comprising lists of 37 jobs, 110 work tasks, 24 spare time activities, and 11 subject areas relevant to the Army. Respondents are asked to indicate whether they would like the jobs, work tasks, spare time activities, and subject areas. Five response options, ranging from "like very much" to "dislike very much," are available for each item. Scores on 22 AVOICE scales are computed on the basis of responses to those items.

Eight AVOICE composites, made up of unit-weighted standard AVOICE scale scores, were used in this study. These composites, which tap broad clusters of interests, were also developed as part of the Experimental Predictor Battery for Project A/Career Force (Campbell & Zook, 1990). The eight AVOICE composites, along with the scales that constitute them, are as follows:

AVOICE Composite	AVOICE Scales
Administrative	Clerical/Administrative
	Warehousing/Shipping
Audiovisual Arts	Aesthetics
	Audiographics
	Drafting
Food Service	Food Service (Employee)
	Food Service (Professional)
Interpersonal	Leadership/Guidance
	Medical Services
Protective Services	Fire Protection
	Law Enforcement

Rugged/Outdoors	Combat Firearms Enthusiast Rugged Individualism
Skilled/Technical	Computers Electronic Communications Mathematics Science/Chemical
Structural/Machines	Electronics Heavy Construction Mechanics Vehicle/Equipment Operator

Job Orientation Blank (JOB). The JOB is a 31-item inventory developed to measure job outcome preferences. This inventory contains a list of job features; respondents are asked to indicate whether or not they would like each feature in their ideal jobs. Five response options, again ranging from "like very much" to "dislike very much," are available for each item. The JOB contains six scales: Job Autonomy, Job Routine, Ambition, Job Pride, Job Security/Comfort, and Serving Others. The first two scales constitute composites by themselves, whereas the standard scores of the latter four scales are summed to form a composite labeled High Expectations.

Thus, first-term attrition was modeled using 26 pre-enlistment predictors, comprising five variables available to the Army at present and 21 non-cognitive variables derived from measures developed during Project A/Career Force. Descriptive statistics for all 26 predictors appear in Table 5.7. The coefficient alpha reliability estimates for the ABLE scales appear in White (1992, p. 26); estimates for the AVOICE and JOB appear in Campbell and Zook (1990).

Table 5.7

Descriptive Statistics for the 26 Pre-Enlistment Predictors by Job Group[a]

Measure	Mean	Std. Dev.
	Job Group C3 (n = 9,690)	
High School Diploma Graduate Status (percent)	0.83	0.37
ASVAB: Quantitative	100.44	13.98
ASVAB: Speed	106.79	11.12
ASVAB: Technical	130.85	18.42
ASVAB: Verbal	102.43	13.40
ABLE: Conscientiousness	36.71	3.98
ABLE: Cooperativeness	44.53	4.86
ABLE: Dominance	27.32	4.51
ABLE: Emotional Stability	40.63	5.34
ABLE: Energy Level	50.90	5.92
ABLE: Internal Control	41.94	4.23
ABLE: Nondelinquency	47.54	5.56
ABLE: Physical Condition	13.82	2.87

(Continued)

Table 5.7 (continued)

Descriptive Statistics for the 26 Pre-Enlistment Predictors by Job Group[a]

Measure	Mean	Std. Dev.
ABLE: Self-Esteem	28.89	3.91
ABLE: Traditional Values	29.04	2.87
ABLE: Work Orientation	45.45	6.07
ABLE: Social Desirability	17.09	3.47
AVOICE: Administrative	98.82	18.45
AVOICE: Audiovisual Arts	149.78	23.38
AVOICE: Food Service	97.94	17.74
AVOICE: Protective Services	101.49	16.61
AVOICE: Rugged/Outdoors	158.33	23.17
AVOICE: Skilled/Technical	200.87	31.47
AVOICE: Social	98.78	16.87
AVOICE: Structural/Machines	205.50	30.37
JOB: Autonomy	49.75	9.86
JOB: High Expectations	196.90	10.01
JOB: Routine	50.81	10.01
Job Group C4 (n = 6,440)		
High School Diploma Graduate Status (percent)	0.87	0.34
ASVAB: Quantitative	106.65	13.05
ASVAB: Speed	108.14	10.79
ASVAB: Technical	139.02	15.85
ASVAB: Verbal	108.53	11.70
ABLE: Conscientiousness	36.99	4.14
ABLE: Cooperativeness	44.80	4.96
ABLE: Dominance	28.00	4.55
ABLE: Emotional Stability	41.17	5.42
ABLE: Energy Level	51.69	6.02
ABLE: Internal Control	42.58	4.22
ABLE: Nondelinquency	47.80	5.50
ABLE: Physical Condition	13.78	2.97
ABLE: Self-Esteem	29.35	3.94
ABLE: Traditional Values	29.25	2.86
ABLE: Work Orientation	46.12	6.18
ABLE: Social Desirability	16.84	3.33
AVOICE: Administrative	96.88	17.63
AVOICE: Audiovisual Arts	149.44	23.00
AVOICE: Food Service	96.61	17.00
AVOICE: Protective Services	103.01	15.88
AVOICE: Rugged/Outdoors	163.87	21.67
AVOICE: Skilled/Technical	200.66	31.02
AVOICE: Social	99.46	16.24
AVOICE: Structural/Machines	204.81	29.76
JOB: Autonomy	50.22	9.61
JOB: High Expectations	196.78	31.05
JOB: Routine	49.04	9.57

(Continued)

Table 5.7 (continued)

Descriptive Statistics for the 26 Pre-Enlistment Predictors by Job Group[a]

Measure	Mean	Std. Dev.
Job Group NC3 (n = 10,563)		
High School Diploma Graduate Status (percent)	0.89	0.32
ASVAB: Quantitative	103.00	13.26
ASVAB: Speed	108.57	11.67
ASVAB: Technical	128.94	18.16
ASVAB: Verbal	105.29	11.29
ABLE: Conscientiousness	37.02	3.99
ABLE: Cooperativeness	44.88	4.83
ABLE: Dominance	27.28	4.58
ABLE: Emotional Stability	40.10	5.54
ABLE: Energy Level	50.72	5.88
ABLE: Internal Control	42.18	4.25
ABLE: Nondelinquency	48.32	5.42
ABLE: Physical Condition	13.18	3.09
ABLE: Self-Esteem	28.81	3.89
ABLE: Traditional Values	29.28	2.74
ABLE: Work Orientation	45.47	6.03
ABLE: Social Desirability	16.76	3.35
AVOICE: Administrative	101.79	18.56
AVOICE: Audiovisual Arts	151.06	23.44
AVOICE: Food Service	101.03	19.38
AVOICE: Protective Services	100.98	18.61
AVOICE: Rugged/Outdoors	141.48	28.59
AVOICE: Skilled/Technical	199.25	31.02
AVOICE: Social	104.67	17.94
AVOICE: Structural/Machines	191.74	33.67
JOB: Autonomy	49.67	10.02
JOB: High Expectations	206.14	29.89
JOB: Routine	50.06	9.69
Job Group NC4 (n = 4,340)		
High School Diploma Graduate Status (percent)	0.91	0.28
ASVAB: Quantitative	104.01	13.18
ASVAB: Speed	107.78	11.41
ASVAB: Technical	133.33	18.16
ASVAB: Verbal	104.34	11.97
ABLE: Conscientiousness	36.73	4.03
ABLE: Cooperativeness	44.40	4.85
ABLE: Dominance	26.82	4.64
ABLE: Emotional Stability	39.79	5.52
ABLE: Energy Level	50.20	5.88
ABLE: Internal Control	41.79	4.22
ABLE: Nondelinquency	47.85	5.40
ABLE: Physical Condition	13.18	3.01

(Continued)

Table 5.7 (continued)

Descriptive Statistics for the 26 Pre-Enlistment Predictors by Job Group[a]

Measure	Mean	Std. Dev.
ABLE: Self-Esteem	28.53	3.91
ABLE: Traditional Values	29.19	2.74
ABLE: Work Orientation	45.21	6.11
ABLE: Social Desirability	16.68	3.37
AVOICE: Administrative	102.80	19.06
AVOICE: Audiovisual Arts	149.61	23.13
AVOICE: Food Service	100.28	19.39
AVOICE: Protective Services	97.51	17.78
AVOICE: Rugged/Outdoors	145.72	26.08
AVOICE: Skilled/Technical	201.59	31.08
AVOICE: Social	97.56	17.36
AVOICE: Structural/Machines	207.48	30.95
JOB: Autonomy	49.71	9.98
JOB: High Expectations	200.25	30.48
JOB: Routine	50.57	9.46

[a] Job Groups: C3, Combat, 3-year enlistment; C4, Combat, 4-year enlistment; NC3, Non-Combat, 3-year enlistment; NC4, Non-Combat, 4-year enlistment.

Analysis

Testing the proportionality assumption. As described in equation 2, Cox's model allows the baseline hazard function, $h_0(t)$, to take any form, with the ratio of the hazards for any two individuals or groups assumed to be a constant value. Before estimating the proportional hazards regression equations, an empirical test of the proportionality assumption was conducted for each MOS within each of the four job groups by examining the significance of the parameter for a single MOS dummy variable (D_{MOS}) interacting with time. Following the general form given in equation 4, the following models were calculated:

$$\begin{aligned} \ln[h(t)] &= \ln[h_0(t)] + \beta_1 D_{MOS} \\ \ln[h(t)] &= \ln[h_0(t)] + \beta_1 D_{MOS} + \beta_2 X \end{aligned} \tag{8}$$

where $h(t)$ is the hazard rate, $h_0(t)$ is the baseline hazard for the jobs not modeled by D_{MOS}, and X is the interaction between D_{MOS} and the log of the event times. If β_2 is significant, then the ratio of the hazard plots is not constant, indicating that the hazard function for the MOS modeled by D_{MOS} is not proportional to the aggregate hazard for the several other MOS in the job group. In all instances, the β_2 parameter was significant at $p < .001$, indicating the proportionality assumption is untenable. The baseline survivor

and hazard functions for one MOS from each of the four job groups are given in Figures 5.3 and 5.4, respectively.[3]

Fitting the proportional hazards models. Because the hazards proved not to be proportional, a stratified analysis was conducted where "MOS" was the stratifying variable. Two types of proportional hazard analyses were performed for each of the four job groups.

The first set of analyses involved predictor blocks that were entered hierarchically. The first block contained the four ASVAB composites and the dummy variable HSDG denoting high school diploma graduate status (HSDG=1 if a diploma graduate, zero otherwise). This block comprises the "base" variables -- the pre-enlistment variables currently available to the Army, against which incremental prediction was assessed. As such, block one was included in all models. Blocks two through four contained the ABLE, AVOICE, and JOB scores, respectively. The incremental fit afforded by each block was assessed relative to block one. Finally, a model was estimated containing all four blocks of predictors.

The second type of analysis used a best subset selection algorithm in the SAS procedure PHREG (SAS Institute, 1992), in which the best j equations containing k specified predictors are given. The model considered to be best for a given number of predictors is the one yielding the highest global score chi-squared statistic. Here, the best j = 3 equations containing k = 1, . . . , 26 predictors were obtained. The value for k is a function of the number of predictions that produce statistically significant increments in model fit. Likelihood ratio tests were calculated to evaluate the point at which additional predictors failed to increase the fit of the model significantly. Nested equations were selected from the best equations and re-evaluated in PHREG to obtain the value of -2 Log L for each model. Stringent p-values were selected, due to the large samples and the dependence of χ^2 on sample size. The effect of each variable, conditional on all other variables in the model, was also taken into account when deriving the "best" equation for each of the four groups.

---

[3] When the values of the predictor variables do not vary across time, the survivor function can be written

$$S(t;\ x_n) \ = \ [\, S_0(t)\,]^{\exp(\Sigma \boldsymbol{x}_n' \boldsymbol{\beta})}$$

where X_n is the vector of predictor variables for the n^{th} person, and $S_0(t)$ is the baseline survivor function,

$$S_0(t) \ = \ \exp(-\int_0^t h_0(u)\, du) \quad .$$

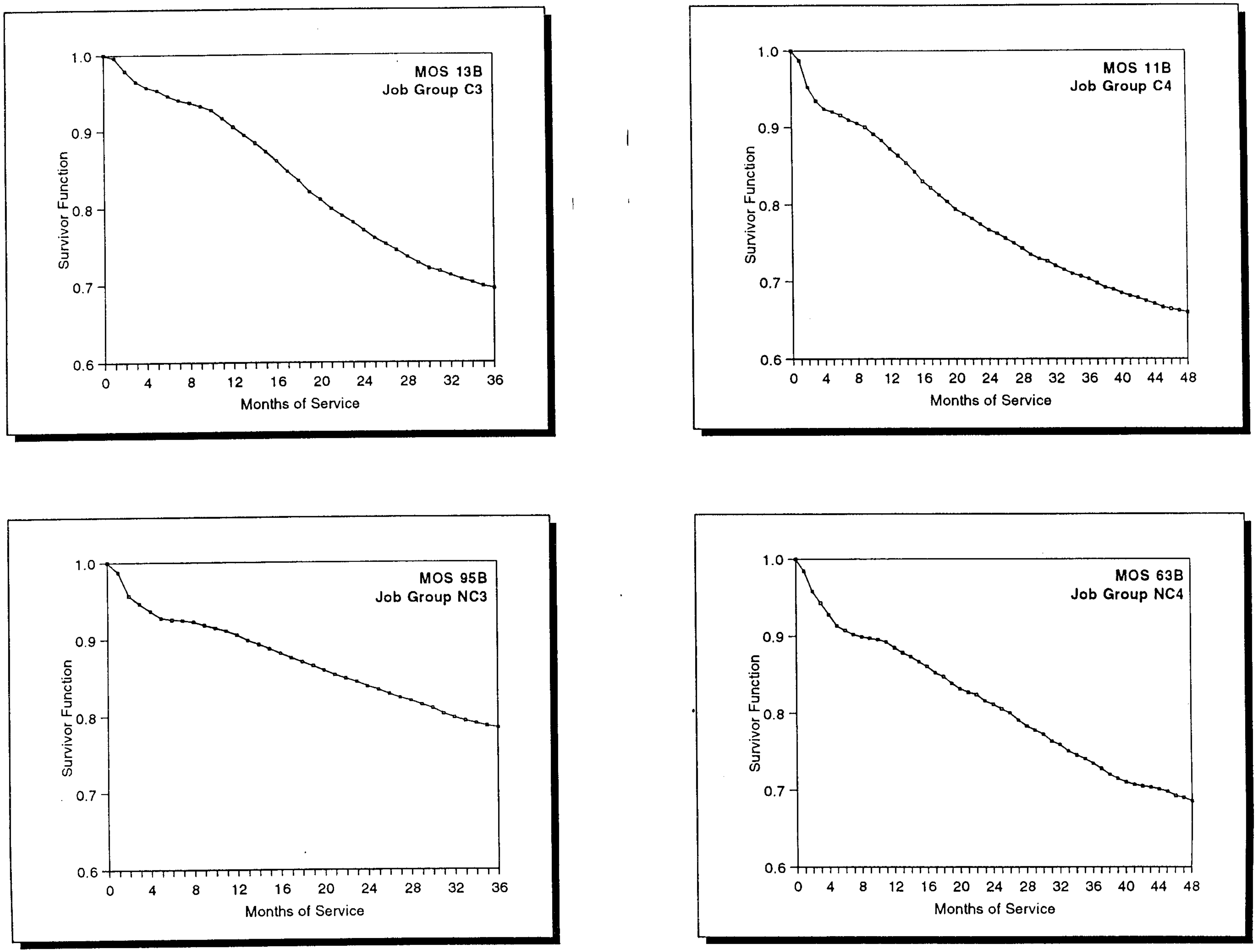


Figure 5.3. Baseline survivor functions for one MOS from each of the four job groups.

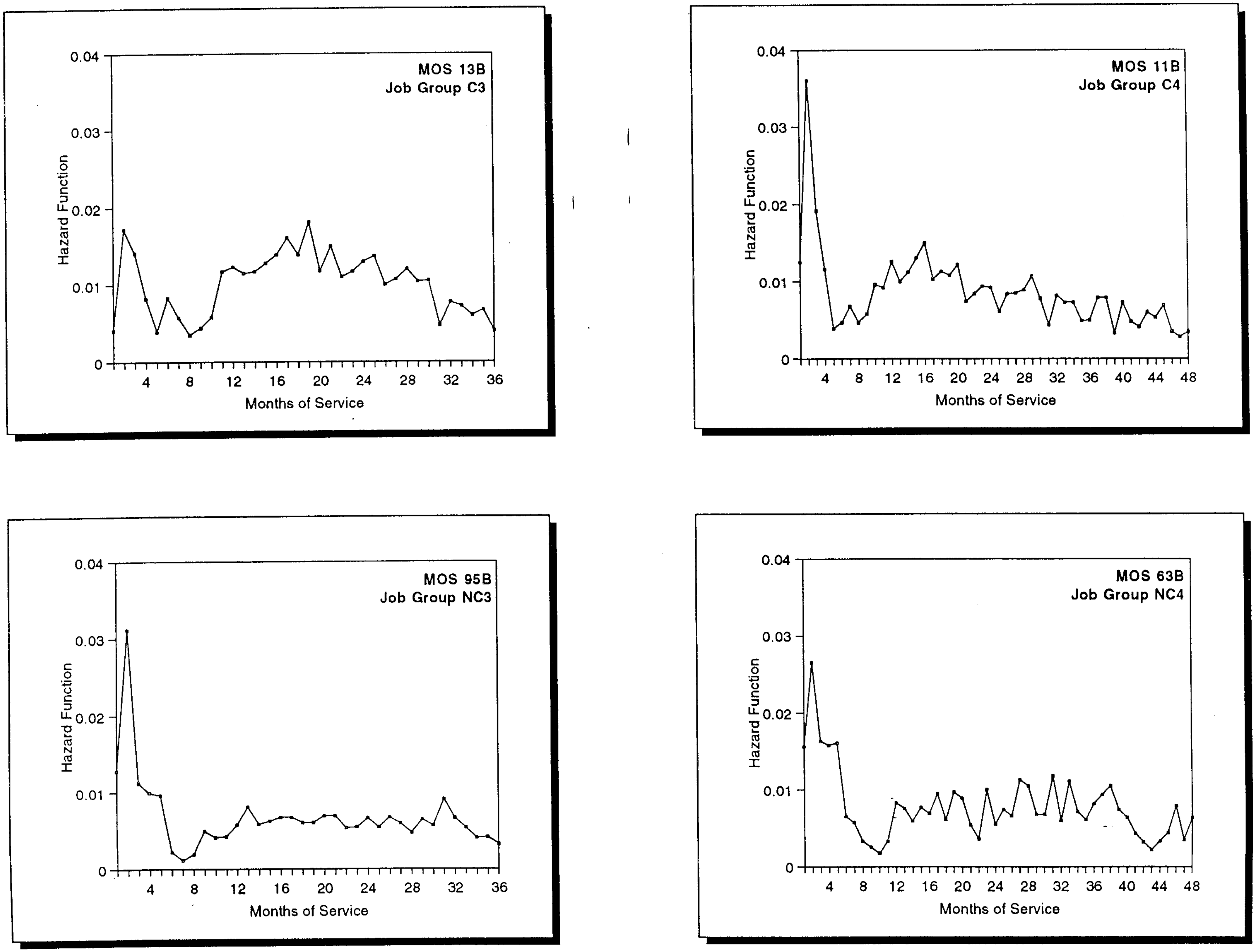


Figure 5.4. Baseline hazard functions for one MOS from each of the four job groups.

Results

Predictor-Block Models

The results for the five attrition models involving predictor blocks for each of the four job groups are given in Table 5.8. (The parameters for each of the models are given in an Appendix available from the first author.) The top half of the table reports the values of -2 Log L for the five models. (For stratified analyses, the likelihood, L, maximized for a job group is the product of each MOS-specific likelihood.) The bottom half of the table contains likelihood ratio test results (based upon the values in the top half of the table) for specific model comparisons.

Consider job group C3 as an example. The value of -2 Log L for Model 1 (the base variables) is 434.02 with 5 degrees of freedom ($p < .0001$). Clearly, the base variables provide significant prediction of first-term attrition for soldiers with 3-year terms of enlistment in Combat jobs. Model 2 adds the 12 ABLE measures to the base variables. The value of -2 Log L increases to 653.08 (again, $p < .0001$). To examine whether this represents a statistically significant increase in the fit of the model to the data, we calculate the difference between -2 Log L for the two models. Hence,

$$653.08 - 434.02 = 219.06$$

with (17 - 5) = 12 degrees of freedom. The critical value for χ^2 at $p < .001$ is 32.91, indicating that the ABLE provides a significant increase in model fit. The other values in the bottom half of Table 5.8 were calculated similarly.

The results in Table 5.8 indicate the following:

- The base variables are significantly related to first-term attrition.
- The ABLE provides significant incremental fit to the base model for all four job groups.
- The AVOICE significantly increases the fit of the base model for 3-year enlistments but not 4-year enlistments. Although the sample sizes differ by a factor of nearly two to one (3-year N = 20,252 and 4-year N = 10,780), this is probably not a power issue, given the absolute size of each group.
- Except for NC3, the JOB does not provide a statistically significant increase in fit over the base model.
- Addition of the AVOICE and JOB to the ABLE model increases the fit of the model to the data for job groups C3 and NC3 (Model 2 vs. Model 5); again, this is probably not a power issue.

Thus, the non-cognitive measures improve prediction of first-term attrition over and above current pre-enlistment information.

Table 5.8

Fit Statistics for the Five Models Involving Predictor Blocks for Predicting First-Term Attrition by Job Group

Values of -2 Log L:

		Job Group[a]			
Model	df	C3	C4	NC3	NC4
(1) Base[b]	5	434.02	188.48	329.72	102.30
(2) Base + ABLE	17	653.08	330.67	649.52	272.58
(3) Base + AVOICE	13	478.96	207.32	352.78	115.67
(4) Base + JOB	8	444.08	198.97	356.95	105.05
(5) Base + ALL	28	693.84	351.96	690.04	284.40

Values of -2 Log (L2 - L1):

		Job Group			
Model Comparisons	df	C3	C4	NC3	NC4
1 vs. 2	12	219.06	142.19	319.80	170.28
1 vs. 3	8	44.94	18.84[ns]	23.06*	13.37[ns]
1 vs. 4	3	10.06[ns]	10.49[ns]	27.23	2.75[ns]
1 vs. 5	23	259.82	163.48	360.32	182.10
2 vs. 5	11	40.76	21.29[ns]	40.52	11.81[ns]
3 vs. 5	15	214.88	144.64	337.26	168.73
4 vs. 5	20	249.76	152.99	333.09	179.35

All values significant at $p < .001$ except * $p < .01$ and [ns] non-significant.

[a] C3, Combat, 3-year enlistment; C4, Combat, 4-year enlistment; NC3, Non-Combat, 3-year enlistment; NC4, Non-Combat, 4-year enlistment.

[b] Base = Four ASVAB composites plus high school diploma graduate status (HSDG).

Best Subset Selection

The "variable traces" of the nested models for the best subset selection analyses are given in Tables 5.9 through 5.12, along with the associated -2 Log L values and differences between them. All likelihood ratio tests were evaluated as χ^2 with one degree of freedom. The tests for groups C3 and NC3 were evaluated relative to a critical value of 10.8 ($p < .001$); tests for groups C4 and NC4 were evaluated relative to a critical value of 7.9 ($p < .005$). The traces end at the point of the last significant increase in model fit. All models begin with k = 3 predictors.

As shown in Tables 5.9 through 5.12, the number of significant predictors ranges from seven to eleven. No AVOICE or JOB scales appear in any of the equations. Although the equations are slightly different across the job groups, six variables appear in each best model: high school diploma graduate status, and five ABLE scales (Nondelinquency, Dominance, Physical Condition, Self-Esteem, and Social Desirability). In addition, the ASVAB Quantitative composite appears in all but one best equation (that for NC4). Indeed, these seven variables constitute the best equation for both Combat groups.

Table 5.9

Variable Traces and Fit Statistics of the Nested Attrition Prediction Models for the Best Subset Selection Analyses: Combat, 3-Year Job Group (C3)

Model	-2 Log L	χ^2 (df = 1)
HSDG, QUANT,[a] Nond[b]	41,353.64	
+ Dominance	41,328.84	24.80
+ Physical Condition	41,310.58	18.26
+ Social Desirability	41,290.80	19.78
+ Self-Esteem	41,277.66	13.14

Significant χ^2 Values, df = 1: $p < .001$ = 10.83. n = 9,689

[a] ASVAB Quantitative composite

[b] ABLE Nondelinquency scale

Table 5.10

Variable Traces and Fit Statistics of the Nested Attrition Prediction Models for the Best Subset Selection Analyses: Combat, 4-Year Job Group (C4)

Model	-2 Log L	χ^2 (df = 1)
HSDG, QUANT,[a] Nond[b]	31,667.26	
+ Dominance	31,651.36	15.90
+ Self-Esteem	31,631.08	20.28
+ Social Desirability	31,620.74	10.34
+ Physical Condition	31,610.32	10.42

Significant χ^2 Values, df = 1: $p < .005$ = 7.88. n = 6,440

[a] ASVAB Quantitative composite

[b] ABLE Nondelinquency scale

Table 5.11

Variable Traces and Fit Statistics of the Nested Attrition Prediction Models for the Best Subset Selection Analyses: Non-Combat, 3-Year Job Group (NC3)

Model	-2 Log L	χ^2 (df = 1)
HSDG, QUANT,[a] Cond[b]	39,482.91	
+ Nondelinquency	39,411.01	71.90
+ TECHNICAL	39,387.20	23.81
+ VERBAL	39,358.51	28.69
+ Social Desirability	39,325.83	32.68
+ Emotional Stability	39,302.77	23.06
+ Dominance	39,269.28	33.49
+ Energy Level	39,253.28	16.00
+ Self-Esteem	39,236.97	16.31

Significant χ^2 Values, df = 1: $p < .001$ = 10.83. n = 10,563

[a] ASVAB Quantitative composite

[b] ABLE Physical Condition scale

Table 5.12

Variable Traces and Fit Statistics of the Nested Attrition Prediction Models for the Best Subset Selection Analyses: Non-Combat, 4-Year Job Group (NC4)

Model	-2 Log L	χ^2 (df = 1)
HSDG, Nond,[a] Cond[b]	16,564.75	
+ Dominance	16,534.93	29.82
+ TECHNICAL	16,518.38	16.55
+ VERBAL	16,506.40	11.98
+ Social Desirability	16,491.86	14.54
+ Self-Esteem	16,482.54	9.32

Significant χ^2 Values, df = 1: $p < .005$ = 7.88. n = 4,340

[a] ABLE Nondelinquency scale

[b] ABLE Physical Condition scale

The parameter estimates for each best equation and their standard errors and p-values are provided in Tables 5.13 through 5.16. To help determine the conditional impact of each variable on the hazard of first-term attrition, values of the risk ratio are also provided. The risk ratio is the exponentiated value of the regression coefficient (i.e., e^b). These values give some insight into the effect of the independent variables on the criterion, subject to the usual caveats applied when determining the effect of a predictor on a criterion in conventional multiple regression analyses based on the value of the regression coefficients (e.g., Darlington, 1972). A value less than one indicates a decrease in the hazard rate (the regression coefficient is negative); a value greater than one indicates an increase in the hazard rate.

Table 5.13

Parameters and Risk Ratios for the Best Attrition Prediction Equation: Combat, 3-Year Job Group (C3)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio One-Unit Increase	Risk Ratio One-SD Increase
HSDG	-.705	.045	.0001	.494	--
QUANTITATIVE	-.013	.002	.0001	.987	.836
Nondelinquency	-.047	.004	.0001	.954	.769
Dominance	.039	.006	.0001	1.040	1.194
Phys. Cond.	-.023	.008	.0014	.976	.933
Social Desir.	.030	.007	.0001	1.031	1.113
Self-Esteem	-.026	.007	.0003	.974	.903

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	41,907.32	41,277.66	629.66 (p = .0001)

n = 9,689

The parameters in Tables 5.13 through 5.16 were calculated in the raw metric of the predictors. Hence, they are analogous to raw coefficients. As with raw coefficients in conventional regression, the risk ratio indicates the effect on the hazard associated with a one-unit increase in the predictor variable. For example, in the best equation for job group C3, the risk ratio for Social Desirability of 1.031 signifies that a one-unit increase in the

Table 5.14

Parameters and Risk Ratios for the Best Attrition Prediction Equation: Combat, 4-Year Job Group (C4)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio One-Unit Increase	Risk Ratio One-SD Increase
HSDG	-.566	.057	.0001	.568	--
QUANTITATIVE	-.008	.002	.0001	.992	.902
Nondelinquency	-.040	.005	.0001	.961	.803
Dominance	.039	.007	.0001	1.040	1.196
Self-Esteem	-.029	.008	.0004	.971	.891
Social Desir.	.026	.008	.0008	1.026	1.090
Phys. Cond.	-.027	.008	.0012	.973	.923

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	31,927.11	31,610.32	316.79 (p = .0001)

n = 6,440

score on this ABLE scale increases the hazard rate for attrition by 3.1 percent.[4]

If one wishes to determine the impact on the hazard rate of an increase of x units (or decrease of -x units), one simply raises the risk ratio to the power of x (or -x). For example, to determine the increase on the hazard rate for a five-unit increase in the Social Desirability score, one computes $(1.031)^5 = 1.165$. Thus, a five-unit increase in the Social Desirability score increases the hazard of first-term attrition by 16.5 percent.

This same approach allows one to determine the effect of an increase of one standard deviation (σ) in the measures, which is equivalent to obtaining standardized regression coefficients. The metrics of the predictors vary to a

[4] In the special case of dichotomous (i.e., dummy) variables, the risk ratio describes the difference in the hazard rates for the two groups. For example, the risk ratio for HSDG of .494 suggests that the hazard rate for high school graduates is 49.4 percent of that for non-graduates. Equivalently, the hazard rate for non-graduates is (1/0.494) = 2.024, or 202.4 percent, or slightly over twice that of high school graduates.

Table 5.15

Parameters and Risk Ratios for the Best Attrition Prediction Equation: Non-Combat, 3-Year Job Group (NC3)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio One-Unit Increase	Risk Ratio One-SD Increase
HSDG	-.712	.054	.0001	.491	--
QUANTITATIVE	-.012	.002	.0001	.989	.858
Phys. Cond.	-.066	.007	.0001	.936	.816
Nondelinquency	-.039	.004	.0001	.962	.808
TECHNICAL	-.010	.001	.0001	.990	.835
VERBAL	.014	.002	.0001	1.014	1.165
Social Desir.	-.033	.005	.0001	.968	1.128
Emotional Stab.	.036	.007	.0001	1.037	.833
Dominance	.032	.006	.0001	1.032	1.157
Energy Level	.026	.005	.0001	1.027	1.169
Self-Esteem	-.030	.008	.0001	.970	.887

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	39,877.43	39,236.97	640.46 (p = .0001)

n = 10,563

considerable degree, even within the same instrument. Hence, risk ratios based upon one-unit increases could be of little interest. To facilitate metric-free comparisons, Tables 5.13 through 5.16 also contain the values of the risk ratios raised to the power of σ.

The results indicate that the largest conditional effects are quite similar across the four job groups, with the largest effects provided by the ABLE scales Nondelinquency and Dominance. A higher Nondelinquency score translates into a decreased hazard for attrition, whereas the opposite is true for Dominance. For example, a one-unit increase in the Nondelinquency score for Combat soldiers with three-year enlistments decreases the hazard for attrition to 95.4 percent of its current level. By comparison, a one-standard deviation increase in the nondelinquency score results in a 76.9 percent decrease in the hazard. Equivalently (and perhaps easier to picture), a

Table 5.16

Parameters and Risk Ratios for the Best Attrition Prediction Equation: Non-Combat, 4-Year Job Group (NC4)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio: One-Unit Increase	Risk Ratio: One-SD Increase
HSDG	-.568	.084	.0001	.567	--
Nondelinquency	-.046	.006	.0001	.955	.779
Phys. Cond.	-.067	.010	.0001	.935	.816
Dominance	.046	.008	.0001	1.047	1.239
TECHNICAL	-.011	.002	.0001	.989	.824
VERBAL	.013	.003	.0001	1.013	1.163
Social Desir.	.038	.009	.0001	1.038	1.136
Self-Esteem	-.031	.010	.0022	.970	.887

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	16,726.49	16,482.54	243.95 (p = .0001)

n = 4,340

one-unit decrease in the Nondelinquency score yields risk ratios of 1/.954 = 1.048, signifying an increase in the hazard of 4.8 percent. A one-standard deviation decrease would yield ratios of 1/.769 = 1.300, signifying a 30 percent increase in the hazard.[5]

Summary

The results of the block analyses were reasonably consistent across the four job/enlistment term groups. The best subset analyses yielded remarkably similar results across the four groups, identifying six variables (high school diploma graduate status, Nondelinquency, Dominance, Physical Condition, Self-Esteem, and Social Desirability) that appear in all four "best" equations and one (the Quantitative composite from the ASVAB) that appears in three of the four. These seven variables were combined to yield a single composite for predicting first-term Army attrition.

[5] The distribution of Nondelinquency is negatively skewed, with a mean for C3 of 47.5, a standard deviation of 5.6, and a maximum value of 60. Thus, full standard deviation increases in the Nondelinquency scale score might be difficult to attain for a large number of individuals.

Analyses of the Attrition Composite

Calculating regression weights. Because the seven variables chosen for the attrition composite represent the best equation for the two Combat job groups, regression weights are available for these variables from the best subset selection analyses for the Combat groups (see Tables 5.13 and 5.14). However, for the Non-Combat job groups, regression weights for the seven predictors were not available from the previous analyses. Thus, the first step in analyzing the attrition composite was to estimate a proportional hazards regression model for job groups NC3 and NC4 using these seven variables. The resulting equations are given in Tables 5.17 and 5.18. Although not selected as the best equation for the two Non-Combat job groups, the seven variables clearly evidence a significant relationship with first-term attrition in these groups, as well.

Table 5.17

Parameters and Risk Ratios for the Attrition Composite: Non-Combat, 3-Year Job Group (NC3)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio: One-Unit Increase	Risk Ratio: One-SD Increase
HSDG	-.674	.052	.0001	.510	--
Nondelinquency	-.037	.004	.0001	.964	.820
Dominance	.032	.006	.0001	1.032	1.155
Phys. Cond.	-.063	.007	.0001	.939	.823
Self-Esteem	-.036	.007	.0001	.964	.867
Social Desir.	.035	.007	.0001	1.035	1.122
QUANTITATIVE	-.013	.002	.0001	.987	.841

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	39,877.43	39,345.51	531.92 (p=.0001)

n = 10,563

Determining effects of the composite at various cut scores. One might ask what increase in survivability would be attained if the seven-variable attrition composite were used to screen recruits. The appropriate baseline for comparison would be the attrition behavior of the present sample of first-term soldiers.

Table 5.18

Parameters and Risk Ratios for the Attrition Composite: Non-Combat, 4-Year Job Group (NC4)

Variable	Regression Coefficient (b)	Standard Error ($\sigma_{se(b)}$)	p	Risk Ratio One-Unit Increase	Risk Ratio One-SD Increase
HSDG	-.572	.083	.0001	.564	--
Nondelinquency	-.044	.006	.0001	.957	.788
Dominance	.048	.008	.0001	1.049	1.249
Phys. Cond.	-.070	.010	.0001	.932	.809
Self-Esteem	-.030	.010	.0030	.971	.891
Social Desir.	.033	.009	.0003	1.034	1.119
QUANTITATIVE	-.006	.002	.0075	.994	.924

	Without Covariates	With Covariates	χ^2 (df = 7)
-2 Log L	16,726.49	16,504.01	222.48 (p = .0001)

n = 4,340

To examine this question, each person was scored on the appropriate optimally weighted attrition composite. The same seven variables appear in each composite, but each variable is weighted by the regression coefficients given in Tables 5.13, 5.14, 5.17, and 5.18. The distribution of scores on each composite was analyzed using the SAS procedure UNIVARIATE. From these analyses, the scores representing the 10th, 25th, and 33rd percentiles on the attrition composite were identified. Although the first two cut scores were chosen arbitrarily, the 33rd percentile was selected to represent the forecasted maximum reduction expected from the force drawdown.[6] The composite distribution was truncated at these three cut points.

[6] Janice Laurence, personal communication, 7 October 1993.

Descriptive statistics were calculated within job group for the attrition composite and the seven component variables for those individuals scoring above each cut score (see Tables 5.19 and 5.20). Survival curves were also obtained. From the descriptive statistics, the relative increase in the component score with increasing truncation of the composite distribution was represented by an effect size,

$$Effect\ size = \frac{m_t - m_b}{\sigma_b}$$

where m_t and m_b are the means of the component in the truncated and baseline groups, respectively, and σ_b is the baseline group standard deviation. Survival curves were plotted to indicate the increase in survival expected for those scoring above the attrition composite cut scores (see Figures 5.5 through 5.8).

Table 5.19

Descriptive Statistics for the Attrition Composite at Various Truncation Points by Job Group

Job Group	N	Mean	SD	Minimum	Maximum
C3 - Baseline	9,690	3.61	.46	1.52	4.87
C3 - 10% Cut	8,721	3.71	.35	2.96	4.87
C3 - 25% Cut	7,267	3.82	.28	3.34	4.87
C3 - 33% Cut	6,492	3.86	.26	3.47	4.87
C4 - Baseline	6,440	2.95	.37	1.41	3.97
C4 - 10% Cut	5,796	3.03	.28	2.41	3.97
C4 - 25% Cut	4,830	3.12	.21	2.75	3.97
C4 - 33% Cut	4,315	3.15	.20	2.85	3.97
NC3 - Baseline	10,563	4.15	.45	1.90	5.50
NC3 - 10% Cut	9,506	4.24	.34	3.55	5.50
NC3 - 25% Cut	7,922	4.34	.28	3.89	5.50
NC3 - 33% Cut	7,077	4.39	.26	3.99	5.50
NC4 - Baseline	4,340	3.22	.40	.99	4.33
NC4 - 10% Cut	3,906	3.30	.30	2.69	4.33
NC4 - 25% Cut	3,255	3.39	.25	2.99	4.33
NC4 - 33% Cut	2,908	3.43	.23	3.08	4.33

Note: The value of the attrition composite is the sum of its seven variables weighted by their respective regression coefficients (i.e., ΣbX). The magnitude of these values is a function of the coefficients, b, which were obtained when predicting the natural log of the hazard function (i.e., ln [h(t)]) rather than h(t) itself (cf. equations 2 and 6)).

Table 5.20 indicates that, for each job group, the primary means of decreasing attrition is to increase the percentage of high school diploma graduates. Although not particularly surprising, the effect remains striking, with at least 98 percent of high school diploma graduates appearing in the

Table 5.20

Descriptive Statistics at Various Truncation Points for the Attrition Composite and Its Components

Job Group C3 (Combat MOS, 3-Year Term)

	Mean				Standard Deviation				Effect Size		
Variable	No Truncation	Lower 10%	Lower 25%	Lower 33%	No Truncation	Lower 10%	Lower 25%	Lower 33%	Lower 10%	Lower 25%	Lower 33%
Attrition Composite	3.61	3.71	3.82	3.88	0.46	0.35	0.28	0.26	0.22	0.46	0.54
HSDG (percent)	83.13	91.30	98.22	99.20	37.45	28.19	13.21	8.91	0.22	0.40	0.43
Quantitative	100.44	100.66	101.48	102.53	13.98	14.31	14.48	14.41	0.02	0.07	0.15
Dominance	27.32	27.25	27.19	27.15	4.51	4.52	4.54	4.57	-0.02	-0.03	-0.04
Physical Condition	13.82	13.94	14.07	14.15	2.87	2.83	2.79	2.77	0.04	0.09	0.11
Nondelinquency	47.54	48.28	49.03	49.45	5.56	5.05	4.62	4.45	0.13	0.27	0.34
Self Esteem	28.89	29.05	29.25	29.40	3.91	3.85	3.77	3.73	0.04	0.09	0.13
Social Desirability	17.09	17.15	17.16	17.11	3.47	3.49	3.49	3.51	0.02	0.02	0.01
N	9,690	8,721	7,267	6,492							
Composite Cut Score	0	2.957	3.340	3.468							

Job Group C4 (Combat MOS, 4-Year Term)

	Mean				Standard Deviation				Effect Size		
Variable	No Truncation	Lower 10%	Lower 25%	Lower 33%	No Truncation	Lower 10%	Lower 25%	Lower 33%	Lower 10%	Lower 25%	Lower 33%
Attrition Composite	2.95	3.03	3.12	3.15	0.37	0.28	0.21	0.20	0.22	0.46	0.54
HSDG (percent)	87.00	94.94	99.11	99.63	33.63	21.91	9.39	0.06	0.24	0.36	0.38
Quantitative	106.65	107.14	108.34	109.30	13.05	13.19	13.03	12.67	0.04	0.13	0.20
Dominance	28.00	27.95	27.94	27.92	4.55	4.58	4.56	4.56	-0.01	-0.01	-0.02
Physical Condition	13.78	13.91	14.12	14.23	2.97	2.91	2.84	2.81	0.04	0.11	0.15
Nondelinquency	47.80	48.47	49.38	49.83	5.50	5.09	4.48	4.28	0.12	0.29	0.37
Self Esteem	29.35	29.57	29.92	30.09	3.94	3.85	3.68	3.62	0.06	0.14	0.19
Social Desirability	16.84	16.90	16.92	16.94	3.33	3.36	3.36	3.35	0.02	0.02	0.03
N	6,440	5,796	4,830	4,315							
Composite Cut Score	0	2.412	2.749	2.846							

(Continued)

Table 5.20 (continued)

Descriptive Statistics at Various Truncation Points for the Attrition Composite and Its Components

Job Group NC3 (Non-Combat MOS, 3-Year Term)

	Mean				Standard Deviation				Effect Size		
Variable	No Truncation	Lower 10%	Lower 25%	Lower 33%	No Truncation	Lower 10%	Lower 25%	Lower 33%	Lower 10%	Lower 25%	Lower 33%
Attrition Composite	4.15	4.24	4.34	4.39	0.45	0.34	0.28	0.26	0.20	0.42	0.53
HSDG (percent)	88.76	95.40	98.62	99.25	31.58	20.94	11.65	8.62	0.21	0.31	0.33
Quantitative	103.00	103.51	105.02	105.97	13.26	13.38	13.13	12.98	0.04	0.15	0.22
Dominance	27.28	27.35	27.48	27.58	4.58	4.57	4.53	4.53	0.02	0.04	0.06
Physical Condition	13.18	13.41	13.77	13.98	3.09	3.01	2.86	2.78	0.07	0.19	0.26
Nondelinquency	48.32	49.02	49.71	50.03	5.42	4.94	4.60	4.45	0.13	0.26	0.32
Self Esteem	28.81	29.12	29.52	29.73	3.89	3.74	3.55	3.50	0.08	0.18	0.24
Social Desirability	16.76	16.80	16.78	16.77	3.35	3.38	3.34	3.32	0.01	0.01	0.00
N	10,563	9,506	7,922	7,077							
Composite Cut Score	0	3.546	3.885	3.992							

Job Group NC4 (Non-Combat MOS, 4-Year Term)

	Mean				Standard Deviation				Effect Size		
Variable	No Truncation	Lower 10%	Lower 25%	Lower 33%	No Truncation	Lower 10%	Lower 25%	Lower 33%	Lower 10%	Lower 25%	Lower 33%
Attrition Composite	3.22	3.30	3.39	3.43	0.40	0.30	0.25	0.23	0.20	0.42	0.53
HSDG (percent)	91.47	96.59	98.71	99.24	27.93	18.14	11.29	8.67	0.18	0.26	0.28
Quantitative	104.01	104.29	105.09	105.51	13.18	13.28	13.25	13.20	0.02	0.08	0.11
Dominance	26.82	26.74	26.68	26.68	4.64	4.69	4.71	4.71	-0.02	-0.03	-0.03
Physical Condition	13.18	13.43	13.76	13.97	3.01	2.92	2.80	2.74	0.08	0.19	0.26
Nondelinquency	47.85	48.52	49.31	49.70	5.40	4.94	4.55	4.42	0.12	0.27	0.34
Self Esteem	28.53	28.75	29.02	29.22	3.91	3.85	3.79	3.74	0.06	0.13	0.18
Social Desirability	16.68	16.70	16.69	16.73	3.37	3.36	3.35	3.36	0.01	0.00	0.01
N	4,340	3,906	3,255	2,908							
Composite Cut Score	0	2.690	2.987	3.081							

upper 75 percent of the composite distribution. The ABLE scales show relatively modest increases in the truncated groups, with the largest effect sizes (0.32 to 0.37) obtained for Nondelinquency at a truncation point of 33 percent.

Note that although Dominance and Social Desirability both receive positive regression weights in the proportional hazards equations (signifying higher scores are indicative of higher hazards for attrition), the effect sizes for these variables are <u>positive</u> in the majority of job groups (once for Dominance, NC3; and all four times for Social Desirability) and always very small (ranging from 0.00 to 0.06).

Figures 5.5 through 5.8 demonstrate the effects of the truncation on survivability. As a reference point, these figures contain lines indicating the times at which each group would have experienced attrition of 20 percent. For all four plots, the truncation point of 10 percent translates to an increase in the expected survivability of three to four months, whereas the increase evidenced at 25 percent is seven to eight months. The maximal truncation of 33 percent affords an increase in survivability of between nine and ten months, a relatively small increase in survival time over the more modest 25 percent screen.

A second way of looking at the curves is in terms of the percentage of the groups remaining after six months. Six months was chosen because the rate of attrition is highest during the first six months, falling off to a relatively constant rate thereafter. This information for the baseline and three truncation points is given in Table 5.21 for each of the four job groups. Taking a weighted average of the increase in retention, we would expect an average increase in cohort survival across the four job groups of 0.9 percent, 1.8 percent, and 2.1 percent after six months with truncations of 10 percent, 25 percent, and 33 percent, respectively. Similarly, at the two-year point, the expected gains are 1.6 percent, 4.3 percent, and 5.0 percent, respectively.

<u>Increased survivability for diploma graduates and non-graduates</u>. Departing a bit from the attrition composite and focusing on the five ABLE scales that are components, one might ask what increases in survivability obtain for high school diploma graduates and non-graduates who score above the mean on the five ABLE scales relative to graduates and non-graduates who score low on the five ABLE scales. The data in Table 5.20 suggest that most of the reduction in attrition would come from selecting only high school diploma graduates. If so, how much of a contribution do the ABLE scales make to the prediction of attrition, over and above high school diploma graduate status?

To examine this question, the five ABLE scales were standardized within job group, unit weighted (with Dominance and Social Desirability weighted negatively), and summed. This composite score was then standardized within job group. For each job group, high school diploma graduates and non-graduates were identified and split into two groups: those scoring above the mean (ABLE composite > 0 -- High ABLE) and those scoring at or below the mean (ABLE composite ≤ 0 -- Low ABLE).

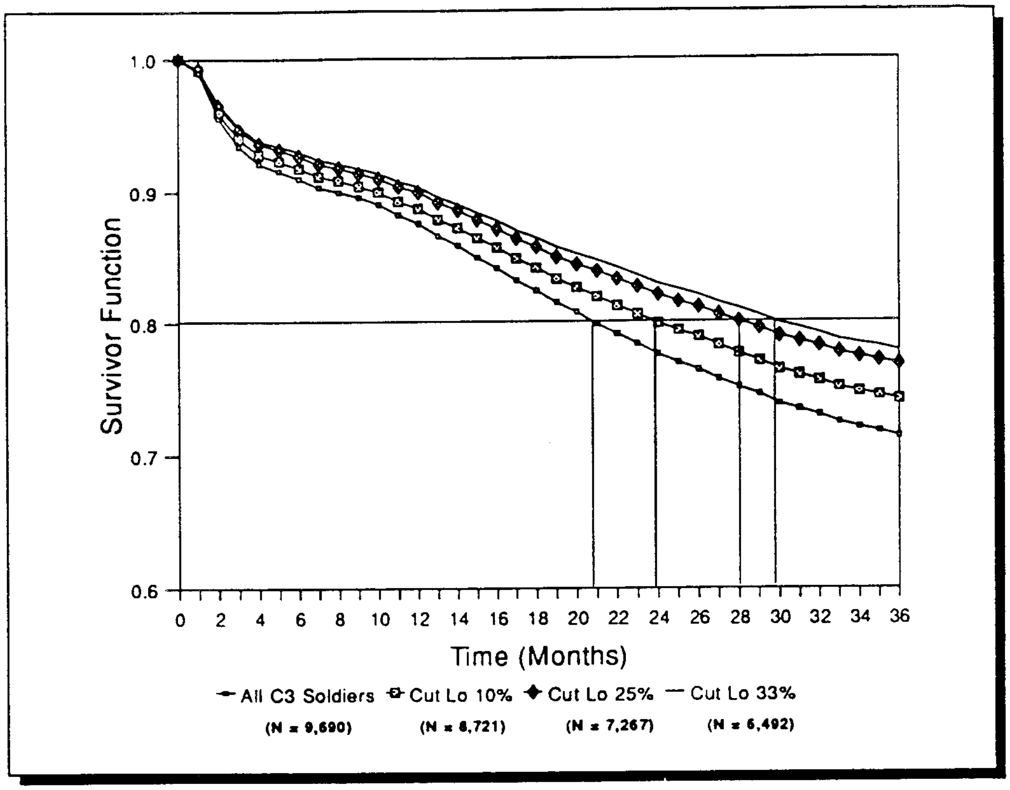


Figure 5.5. Survivor functions for C3 soldiers scoring above various truncation points on the attrition composite.

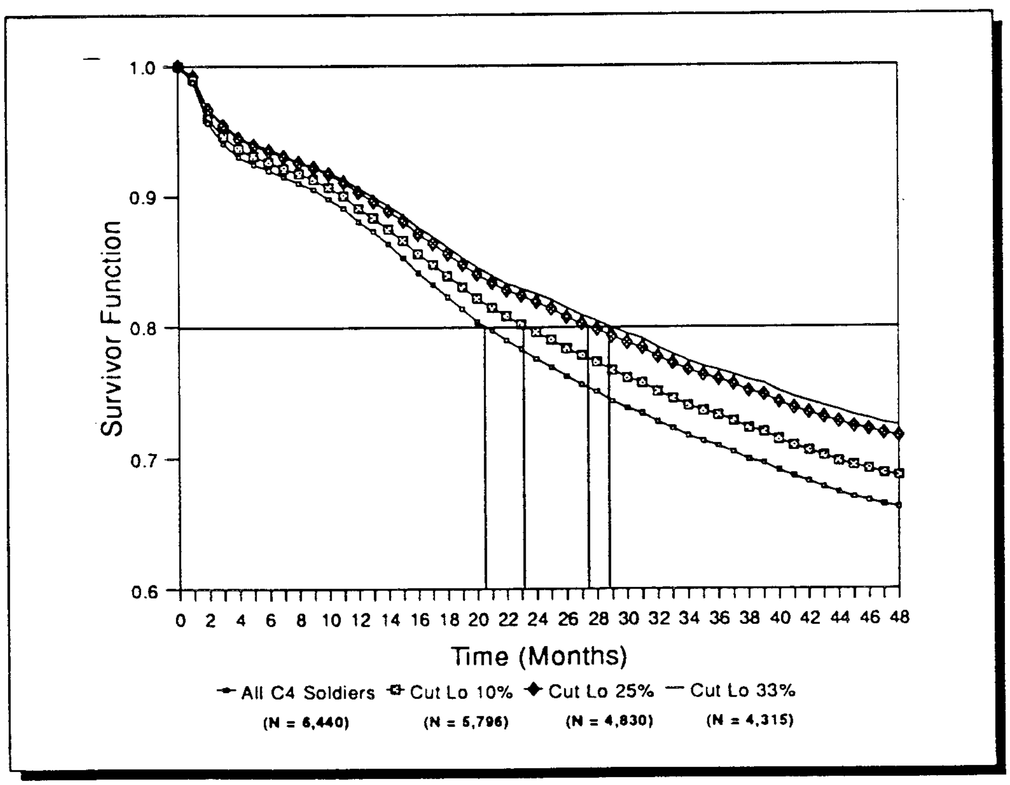


Figure 5.6. Survivor functions for C4 soldiers scoring above various truncation points on the attrition composite.

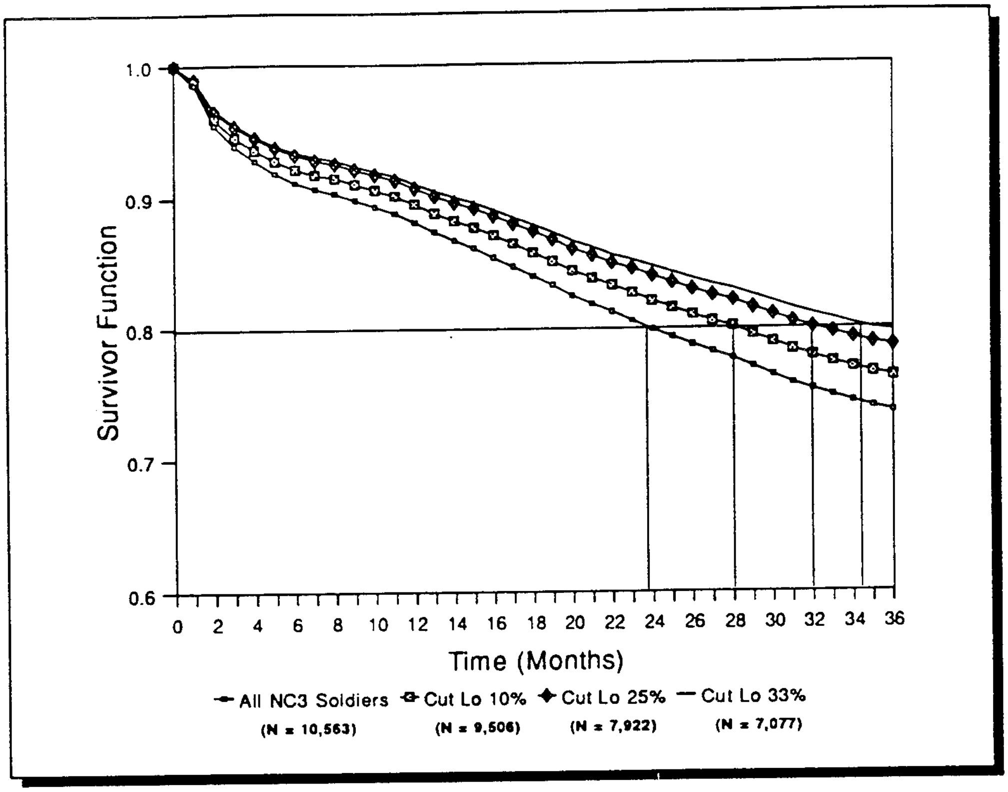


Figure 5.7. Survivor functions for NC3 soldiers scoring above various truncation points on the attrition composite.

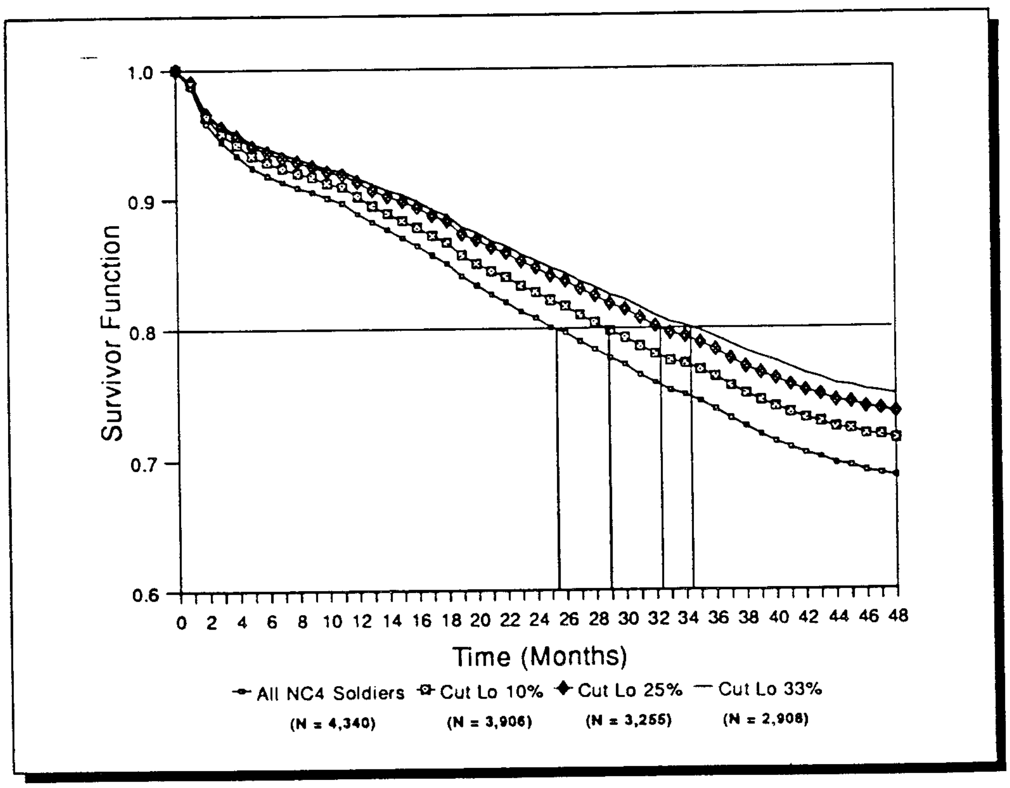


Figure 5.8. Survivor functions for NC4 soldiers scoring above various truncation points on the attrition composite.

Table 5.21

Percent of Cohort Remaining at Various Levels of Truncation on the Attrition Composite

	Job Group			
Truncation	C3	C4	NC3	NC4
Six Months				
No Truncation	91.0	92.0	91.1	91.8
Lower 10%	91.8	92.7	92.2	92.8
Lower 25%	92.7	93.5	93.2	93.6
Lower 33%	93.0	93.7	93.5	93.9
Two Years				
No Truncation	77.6	77.5	79.8	80.8
Lower 10%	80.0	79.5	82.0	82.8
Lower 25%	82.2	81.9	83.9	84.7
Lower 33%	83.0	82.5	84.7	85.2

The number of soldiers in each HSDG/ABLE grouping and the percentage of those who exit early, for each of the job groups, are given in Table 5.22. Note that for all four job groups, more graduates score high than low on the ABLE, whereas the opposite is true for non-graduates.

The effect of high and low ABLE scores on the survivability of graduates and non-graduates is most clearly demonstrated by the survivor functions that are given in Figures 5.9 through 5.12. These figures show the strong effect of high school diploma graduate status, with graduates having higher survival rates than non-graduates. Nevertheless, for both groups of soldiers, those scoring high on the composite formed by the five ABLE scales that are components of the attrition composite have a better rate of survivability than those from the same educational group who score low on the composite.

Perhaps the most striking feature of Figures 5.9 through 5.12 is that there is a greater increase in survivability for those high on ABLE for diploma **graduates** than for non-graduates. That is, although non-graduates who score high on the ABLE demonstrate increased survivability over non-graduates who score low on the ABLE, the difference between survival curves is much larger for graduates. Figures 5.9 through 5.12 indicate that for all job groups except NC3, low-scoring non-graduates require approximately one year for 20 percent of the group to exit prematurely, whereas it takes approximately 14 to 15 months for attrition to reach 20 percent for the high-scoring non-graduates. (The difference for NC3 jobs is approximately seven months. Note that the reference line for both NC3 groups is at .85, representing 15 percent attrition, because the high-scoring graduates do not experience 20 percent attrition during their first term.) By comparison, the difference between high- and low-scoring graduates ranges from 13 to 18 months for the four job groups.

Table 5.22

Number of High School Diploma Graduates and Non-Graduates Scoring Above or Below the Mean on the ABLE Composite

	Job Group	Graduate High ABLE	Graduate Low ABLE	Non-Grad High ABLE	Non-Grad Low ABLE	Total
C3	N	4,402	3,652	604	1,031	9,689
	(Percent)	(45.4)	(37.7)	(6.2)	(10.6)	(100.0)
	Attritions	891	1,062	271	513	2,737
	(Percent)	(20.2)	(29.1)	(44.9)	(49.8)	(28.2)
C4	N	3,033	2,570	308	529	6,440
	(Percent)	(47.1)	(39.9)	(4.8)	(8.2)	(100.0)
	Attritions	779	906	154	285	2,124
	(Percent)	(25.7)	(35.3)	(50.0)	(53.9)	(33.0)
NC3	N	5,040	4,336	417	770	10,563
	(Percent)	(47.7)	(41.0)	(3.9)	(7.3)	(100.0)
	Attritions	937	1,279	165	380	2,761
	(Percent)	(18.6)	(29.5)	(39.6)	(49.4)	(26.1)
NC4	N	2,095	1,875	150	220	4,340
	(Percent)	(48.3)	(43.2)	(3.5)	(5.1)	(100.0)
	Attritions	473	674	76	113	1,336
	(Percent)	(22.6)	(35.9)	(50.7)	(51.4)	(30.8)

This finding is of particular note in light of the research that has attempted to use biodata to select those non-graduates having the best chance of completing their first term of service (McBride, 1993). Figures 5.9 through 5.12 suggest that although the ABLE would be useful for selecting such non-graduates, a greater gain from using the ABLE would be realized if it were used as a screen for identifying the high school diploma graduates who are less likely to complete their first term.

This finding may be due in part to range restriction on the ABLE in the non-high school diploma graduate sample. Because non-graduates are more likely to exit the Army prematurely, they are screened more stringently before they are enlisted than are high school diploma graduates. For example, enlisted non-graduates typically must have high ASVAB scores and pristine moral records. Perhaps requiring a number of compensating characteristics from non-graduates leads to a restriction in individual variability along the dimensions assessed by the ABLE. If so, this would have an adverse effect on the predictive relationship between the ABLE and the attrition behavior of these soldiers.

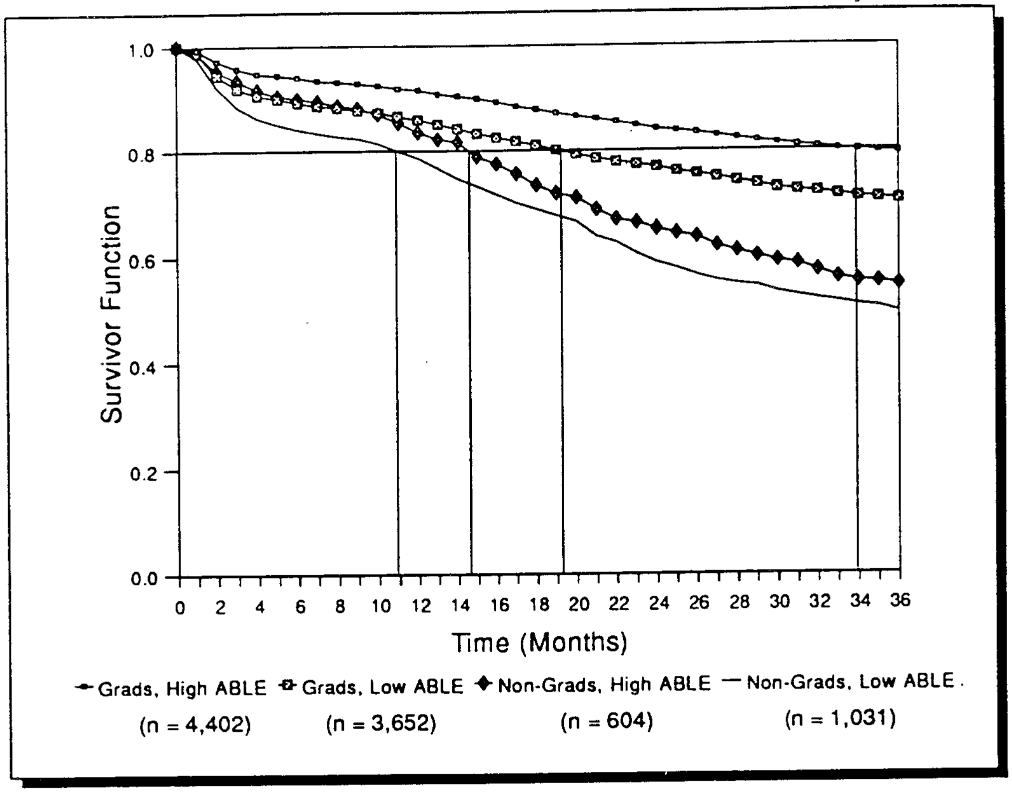


Figure 5.9. Survivor functions for high school diploma graduates and non-graduates scoring above/below the mean of the ABLE composite: Job Group C3.

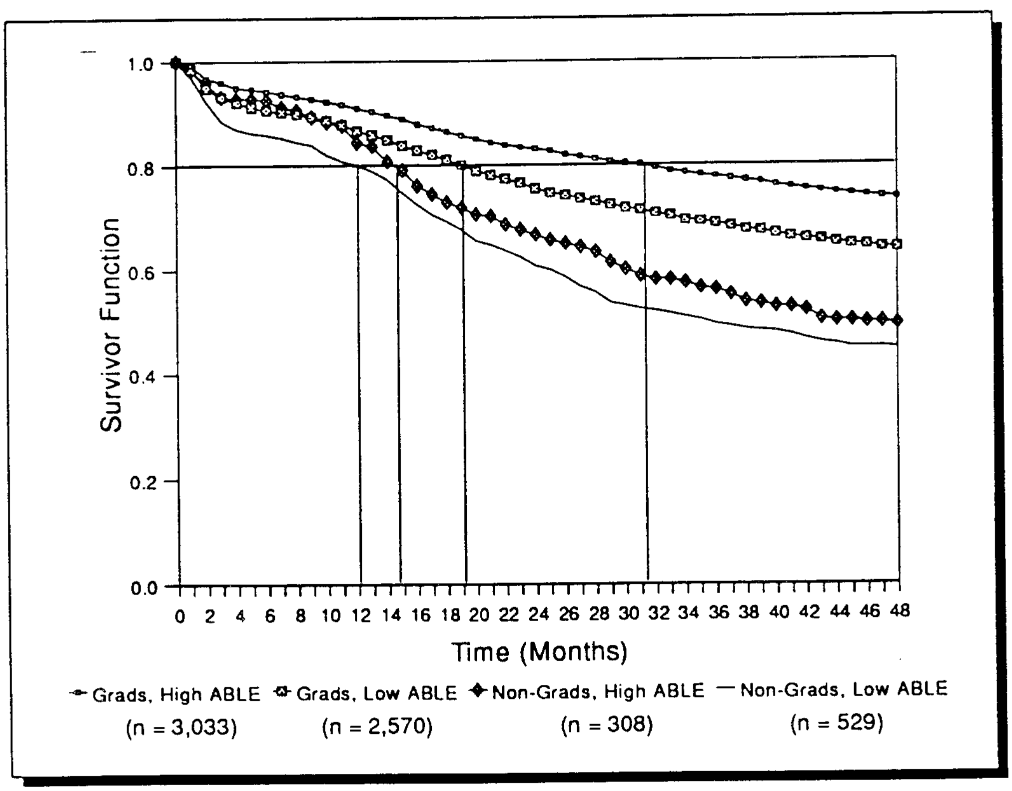


Figure 5.10. Survivor functions for high school diploma graduates and non-graduates scoring above/below the mean of the ABLE composite: Job Group C4.

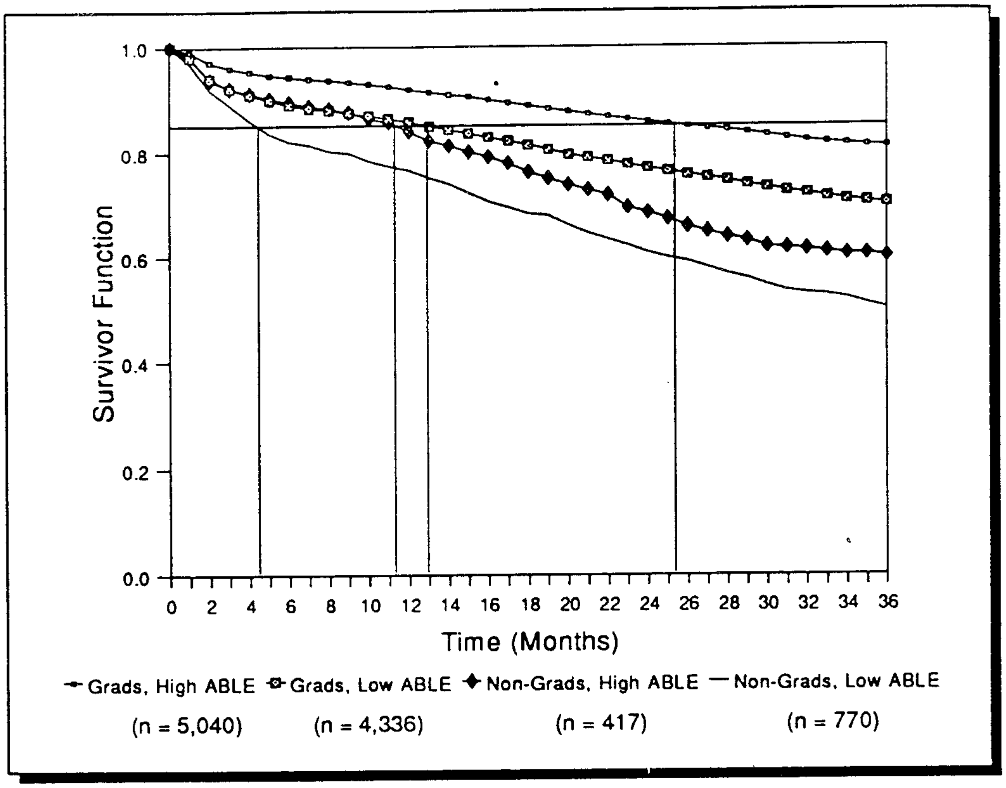


Figure 5.11. Survivor functions for high school diploma graduates and non-graduates scoring above/below the mean of the ABLE composite: Job Group NC3.

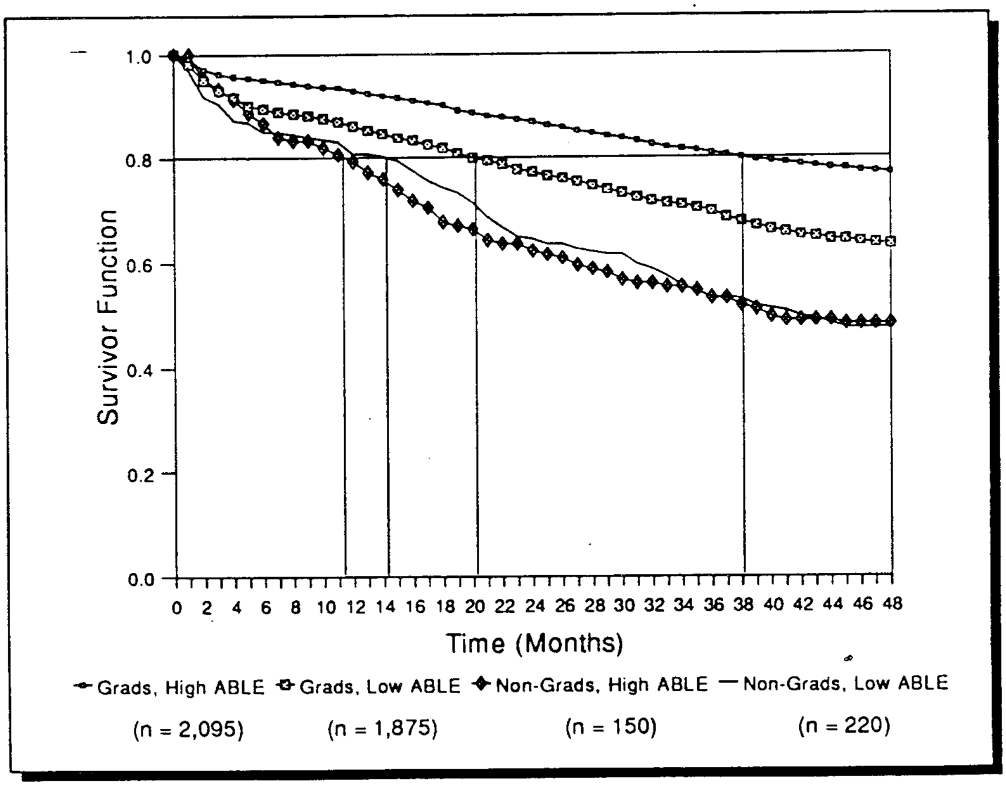


Figure 5.12. Survivor functions for high school diploma graduates and non-graduates scoring above/below the mean of the ABLE composite: Job Group NC4.

Discussion

The results of the event history analysis of first-term attrition indicate that high school diploma graduate status remains a very powerful predictor, with the hazard rate for non-graduates being approximately twice that for graduates. Nevertheless, information provided by non-cognitive measures (specifically, scales from the ABLE) provides significant incremental prediction to pre-enlistment information currently available to the Army. Information provided by the AVOICE and the JOB contributed incremental prediction on occasion, but both were overshadowed when the ABLE was in the equation.

Faking/Coaching

Using biodata and temperament inventories as pre-enlistment selection measures raises questions of the possible effects of coaching and faking on applicant responses. Biodata items tend to be less susceptible to such effects to the extent that they are verifiable. Temperament inventories, however, have been shown to be easily faked (e.g., Hough, Eaton, Dunnette, Kamp, & McCloy, 1990). Hough et al. found, however, that such distortion does not have marked effects on the validity of the instrument. To the extent that this result is generalizable, the concerns about faking might be exaggerated. Note, however, that the respondents in the Hough et al. study were not coached.

The effects of coaching on the validity of the ABLE was investigated by Young, White, and Oppler (1992). Soldiers in this study were asked to complete the ABLE and were assigned to one of three response conditions: Honest, Ad-Lib_Faking (i.e., respondents were to answer the questions in a manner each believed would most impress the Army), and Coached. Soldiers who were coached were given the same instructions given to the Ad-Lib Faking group, but the respondents were then given three practice items and instructed which responses were the most favorable and why. Young et al. examined the effect of the three conditions on the point-biserial correlations between the ABLE scales and a dichotomous criterion variable representing attrition during the first 18 months of service (scored as 1 if attrition occurred, 0 otherwise).

Young et al. found that the coached respondents scored an average of 1.2 and 0.6 standard deviation higher on the ABLE scales than did the Honest and Ad-Lib Faking group respondents. Nevertheless, the correlations across the three conditions were similar in magnitude, with few significant differences being found between the Honest and Ad-lib Faking conditions. In this respect, the results are similar to those reported by Hough et al. (1990) with regard to performance criteria. Virtually all of the point-biserials were negative, indicating higher ABLE scale scores were associated with decreased attrition. The point-biserials for the Coached respondents, however, switched sign, all becoming positive, indicating that higher ABLE scores were associated with increased attrition. Although Young et al. stated that the reasons for this phenomenon were uncertain, the study suggests that the impact of coaching on the validity of the ABLE in an operational setting to screen for first-term attrition could be severe. The effects of coaching, however, might be partially mediated by warnings about scales that can detect faking, score corrections for faking, and so on.

Cross Validity

The use of regression typically raises the question of cross validity. Event history models do not produce a measure similar to the coefficient of determination, obviating the use of shrinkage formulae. Nevertheless, a procedure has been suggested for assessing the fit of a proportional hazards regression model in a second sample (McCloy, 1993). To the extent that the procedure is viable, questions of cross validity may be addressed.

The issue of cross validity, however, has limited application to first-term attrition. Attrition depends in part on organizational policy that in turn depends upon social, economic, and political demands. For example, the current downsizing of the military could result in the Army being a bit more strict in enforcing certain standards for soldiers (e.g., minimum performance standards and physical requirements) than at other times. Optimal equations for predicting attrition that are developed during the observance of a specific organizational policy might cross-validate well as long as the current policy remains the rule. Any change in policy, however, could mean that the equations are no longer optimal. A new analysis would be necessary to develop equations for predicting attrition occurring under the new organizational policy.

Component/Attrition Relationships

The effects of two variables on the hazard of first-term attrition merit special consideration. First is the positive relationship between Dominance and the hazard, indicating that dominant soldiers are more likely to exit prematurely. This relationship might appear counterintuitive, given that dominance could be associated with leadership and a "take-charge" demeanor. Perhaps the finding supports the notion that highly dominant soldiers are recalcitrant and unresponsive to orders, attempting to take charge when they should follow.

Of similar interest is the appearance of the ASVAB Quantitative composite in three of the four best equations. This composite comprises the two mathematics subtests, Arithmetic Reasoning (AR) and Mathematical Knowledge (MK). These two subtests, in turn, appear with the two verbal subtests of Word Knowledge (WK) and Paragraph Comprehension (PC) in the Armed Forces Qualification Test, the military's primary selection measure. Thus, soldiers have already been selected to some degree on the Quantitative score. The AFQT has been shown time and again to have very little relationship to attrition, but the verbal subtests are given twice the weight of the quantitative subtests when calculating the present AFQT score. By removing the Quantitative subtests from the AFQT and using them as a composite, a sizable relationship appears with first-term attrition, notwithstanding the attenuating effects of range restriction.

Indeed, the lack of relationship between the AFQT and first-term attrition could be due to the Quantitative and Verbal subtests working in opposite directions. Whereas higher Quantitative composite scores are associated with a decreased hazard of first-term attrition, the parameters for the Verbal composite from both the block analyses and the best subset analyses are positive. Hence, higher Verbal scores are associated with an <u>increased</u> rate of attrition, with the relationship being negligible for the Combat MOS but relatively strong for the Non-Combat jobs.

The relationships of the Verbal and Quantitative composites might be explained in part by gender differences in attrition rates and scores on the composites. That is, men score higher than women on the Quantitative composite but lower than women on the Verbal composite. In addition, the attrition rate for women is higher than that for men. Thus, the Verbal and Quantitative scores might be serving as proxy variables for a male/female dummy variable. This would explain why the Verbal composite shows no effect in the Combat job groups, where there are no females.

To examine this question, analyses were conducted for the Non-Combat job groups in which a dummy variable for gender was included in the best subset analyses. Although the effect of the dummy variable somewhat mitigated the effect of the Verbal composite, the composite continued to appear in the best prediction equations (i.e., the Verbal composite provided significant incremental fit of the proportional hazards regression model). Thus, the positive relationship between Verbal scores and attrition rate in Non-Combat MOS can not be entirely attributed to a gender effect.

Summary

The current research strongly suggests that the Army can improve the prediction of first-term attrition by gathering biodata/temperament information prior to enlistment and by making better use of the information they currently possess (i.e., ASVAB Quantitative scores). Regarding future research, the hazard functions given in Figure 5.4 indicate that most of the attrition occurs in the first six months of service. The determinants of early-term attrition (e.g., during the first six months) may differ from the determinants of attrition occurring later in the term. Such a hypothesis can be tested using event history analysis. Future research will be directed at this question, as well as further investigation into the issue of cross validity.

Chapter 6

THE ROLE OF JOB SATISFACTION IN PERFORMANCE, ATTRITION, AND REENLISTMENT

Deirdre J. Knapp, Gary W. Carter, Rodney A. McCloy, and Ani S. DiFazio

This chapter begins with a background discussion of the research questions that prompted the inclusion of a job satisfaction measure in the Project A/Career Force research program. It then moves to a description of the development of the Army Job Satisfaction Questionnaire (AJSQ) and its psychometric properties. The remainder of the chapter summarizes Career Force research linking job satisfaction with job performance and two types of military turnover.

BACKGROUND

There were several reasons for adding a job satisfaction instrument to the array of measures administered to Project A/Career Force soldiers. The primary purpose of the measure was to serve as a predictor of turnover -- both non-reenlistment and attrition. The measure was also intended to serve as an additional criterion measure that might be predicted by measures of temperament, interests, and cognitive ability. In this regard, job satisfaction and performance are viewed as parallel criteria. Each is valued and the objective is to increase both job satisfaction and performance. The relationship between facets of job satisfaction and facets of job performance was also a research issue of considerable interest. The specific research questions which were addressed with the Project A/Career Force job satisfaction data are reviewed below.

Research Questions

Satisfaction and Performance

Research Review. Several widely cited literature reviews (e.g., Iaffaldano & Muchinsky, 1985; Podsakoff & Williams, 1986; Vroom, 1964) have presented substantial evidence to show that relationships between job satisfaction and job performance are positive but weak. For example, in a meta-analysis of studies of relationships between job satisfaction and performance, Iaffaldano and Muchinsky found a sample-size weighted mean correlation between job satisfaction and job performance of .146. This mean correlation was based on 217 individual correlations (representing a total sample size of 12,192) between scales tapping a number of different job satisfaction facets and many types of performance measures. When corrected for sampling error and measurement unreliability in both satisfaction and performance measures, this mean correlation rose to .172. Similar findings were reported by Vroom and by Podsakoff and Williams.

Petty, McGee, and Cavender (1984) reported a higher mean correlation (.23), and attributed the difference in mean correlations between that found in their meta-analysis and those found in other quantitative reviews to the inclusion in their review of more studies using professional and managerial samples. The modest size of the mean correlation reported by Iaffaldano and Muchinsky led these researchers to suggest that the satisfaction-performance relation may be an illusory correlation (see Chapman & Chapman, 1969), or a

perceived relation between variables that people believe should be related to each other, but in fact are not.

The reviews cited above summarized data across studies using a diverse array of job performance measures tapping many different aspects of performance. Over the past decade there has been a recognition that job satisfaction can reasonably be expected to be related to certain aspects of performance but not to others. With this has come a renewed appreciation of the importance of carefully considering the relevance of predictor and criterion constructs to each other when designing and interpreting studies.

This recognition led researchers to examine relationships between job satisfaction and extrarole, prosocial behavior. Several researchers (e.g., Bateman & Organ, 1983; Smith, Organ, & Near, 1983) have found moderate correlations between job satisfaction and a cluster of job performance categories that has been labeled organizational citizenship behavior (OCB). Organ (1988) defined OCB as "...behavior that is discretionary, not directly or explicitly recognized by the formal reward system, and that in the aggregate promotes the effective functioning of the organization" (p. 4).

Borman and Motowidlo (1993) discussed contextual performance -- job activities that are not part of the technical core of jobs, or of the "task performance" aspects of jobs, and yet contribute in important ways to organizational effectiveness. Although OCB is similar in many ways to contextual performance, the two classes of behavior do not completely overlap. Organ (1988) stressed that OCB, by definition, "is not an enforceable requirement of the role or the job description" (p. 4) and is not "directly or formally recompensed by the organization's reward system" (p. 5). Contextual performance is a broader construct than OCB, subsuming both extrarole behaviors not explicitly recognized by the formal reward system and behavior that is not part of the technical core of the job and yet is explicitly recognized by the formal reward system (Borman & Motowidlo, 1993).

There are regulations and standard procedures governing virtually every aspect of life among those in the U.S. Army. Moreover, mechanisms exist through which to formally recognize adherence to (or lack of adherence to) these regulations and procedures. For this reason, contextual performance appears to be a more pertinent class of behavior than OCB to study in military organizations.

Hypotheses. It was hypothesized that there would be moderate correlations between job satisfaction and contextual performance indices, but negligible correlations between job satisfaction and scores on "can do" measures of the technical core of the job (i.e., job knowledge and hands-on test scores). Correlations between job satisfaction and individual measures that appear to be influenced by both contextual performance and task performance were expected to be lower than correlations between satisfaction and more pure measures of contextual performance, but higher than correlations between satisfaction and "can do" measures of the technical core of the job.

In addition to these hypotheses being tested at the level of individual measurement instruments, they were tested using the five factor scores from the LVI performance model. In these analyses, the Effort and Leadership and Personal Discipline factor scores were considered to be indices of contextual performance.

Satisfaction and Turnover

Research Review. A great deal of theory development and research has been associated with the link between job attitudes and turnover (see Hom, Prussia, & Griffeth, 1992 and Tett & Meyer, 1993 for recent reviews). Yet it is not clear that research conducted using civilian samples is generalizable to military samples (e.g., Hom, Katerberg, & Hulin, 1979; Knapp, 1993). Furthermore, empirical data from military turnover research are not so plentiful. Indeed, the available military research is rather sparse and has been conducted on diverse samples that may have only limited generalizability to each other (e.g., National Guard enlisted versus Regular Army enlisted).

Very little research has examined the relationship between job satisfaction and military attrition (i.e., separation before the enlistment contract expires). This is understandable given that most attrition occurs early in the enlistment term and is hypothesized to be related to problems of adaptability rather than to job dissatisfaction per se. Although there is more research examining the relationship between job satisfaction and military reenlistment behavior, most studies of reenlistment focus on factors other than job satisfaction (e.g., pay, reenlistment bonuses, spouse support, education).

The amount of relevant military turnover research is also reduced by the fact that turnover research data collected on military personnel who enlisted prior to 1973 (i.e., before the end of the draft) are considered to be of limited generalizability to the All Volunteer Force (Boesel & Johnson, 1984; Etheridge, 1989; Mobley, Hand, Baker, & Meglino, 1979).

Turnover-job satisfaction research conducted on samples of military personnel who enlisted in the mid-1970s or later includes data from two samples of National Guard enlistees collected in the late 1970s (Hom et al., 1979; Hom & Hulin, 1981; Miller, Katerberg, & Hulin, 1979). National Guard personnel work on a part-time basis and their first enlistment term is six years. Satisfaction and reenlistment intention were both strongly correlated with reenlistment behavior in both samples.

Three related studies involved samples of full-time Navy enlisted personnel. Royle and Robertson (1980) appended items on satisfaction and intent to reenlist onto task analysis surveys administered to Navy personnel. Correlations between satisfaction and intent to reenlist ranged from .19 to .39 in the four job samples included in this study. Farkas and Tetrick (1989) reported on a longitudinal study in which job satisfaction and reenlistment intention measures were administered to sailors at three points during their first-term of enlistment. Satisfaction and intention measured after completion of advanced training were significantly correlated with each other and with subsequent reenlistment behavior. Finally, LaRocco, Pugh, and Gunderson (1977) used discriminant analysis to distinguish among three groups of first-tour sailors using several types of variables, including satisfaction and performance. Although the results suggested that both satisfaction and performance contributed to the prediction of group membership, the three groups used by these researchers confounded the criteria of retention and attrition (i.e., one group included attritees and another group included reenlistees).

Youngblood, Mobley, and Meglino (1983) used repeated measures analysis to examine satisfaction, attrition intention, reenlistment intention, and turnover (both attrition and non-reenlistment) data from first-term Marine Corps enlisted personnel. Data were collected at entry, at the end of training, and 18-20 months after entry. Results of the analysis confirmed that the precursors of turnover showed systematic fluctuations throughout the enlistment term. Mobley et al. (1979) reported on an earlier phase of this research which involved analyses of early attrition (within the first 11 weeks) only.

Motowidlo and Lawton (1984) reported the only satisfaction-turnover research found that used Regular Army enlisted personnel. These researchers collected data on expectancies, satisfaction, and reenlistment intentions from two samples of soldiers who were within six months of the end of their first enlistment term. Approximately 15 percent of the 320 soldiers in sample 1 reenlisted and 20 percent of the 299 soldiers in sample 2 reenlisted. The correlation between satisfaction and reenlistment was .20 for sample 1 and .26 for sample 2; correlations between satisfaction and reenlistment intention were .19 and .32, respectively; and correlations between reenlistment intention and reenlistment were .66 and .61.

As indicated by this review, most military turnover-satisfaction research has focused exclusively on reenlistment behavior. Job satisfaction and intention to reenlist have consistently been found to be related to reenlistment behavior. Many questions remain, however. Only one of these studies included a measure of performance (LaRocco et al., 1977). Others have not reported indices of association (or model fit) for the job satisfaction and reenlistment intention measures as predictors of turnover. Clearly, there is more to be learned about affective precursors to turnover in the military arena, even at the most elementary, bivariate levels of analysis.

Hypotheses. It was hypothesized that both job satisfaction and reenlistment intention would be significantly correlated with reenlistment. It was further hypothesized that job satisfaction would provide incremental predictive value for predicting reenlistment over that provided by intention to reenlist. Parallel hypotheses were formulated for the prediction of attrition. That is, it was expected that job satisfaction would be significantly related to attrition even when the intention to attrit was included in the prediction equation. Although it was expected that the models would fit better when the event to be predicted was restricted to voluntary attrition, this hypothesis could not be tested because model fit was confounded with attrition base rate. That is, restricting the event of interest to voluntary attrition reduced the base rate which, in turn, led to decrements in model fit.

Performance and Turnover

Research Review. McEvoy and Cascio (1987) conducted a meta-analysis of the relationship between performance and turnover. The mean correlation, weighted by sample size, was -.28. An appreciable amount of unexplained variance was still apparent after corrections were made for sampling error and measurement unreliability. Moreover, only two of the 24 correlations included in their meta-analysis were from military samples. Specifically, two of the correlations came from the study of Navy regular enlisted personnel by LaRocco et al. (1977) described earlier.

Hypotheses. Because of the paucity of directly related research, job performance indices were included in the reenlistment and attrition analyses on an exploratory basis. That is, no specific hypotheses related to job performance were proposed.

DEVELOPMENT OF THE ARMY JOB SATISFACTION QUESTIONNAIRE (AJSQ)

Given the types of research questions described above, the project needed an instrument with which to measure job satisfaction. It was not feasible to use a standardized instrument (e.g., the Job Descriptive Index, from Smith, Kendall, & Hulin (1969)) for this purpose because of copyright issues. Moreover, although the Army has had a longstanding interest in measuring aspects of satisfaction, motivation, and morale (Borman & Bleda, 1978; Motowidlo et al., 1976), a review of the instruments used in the military literature (e.g., Army Early Career Satisfaction Survey, Army Experience Survey, Commander's Unit Analysis Profile, Cureton's Airman Morale Survey) did not identify instruments that suited the needs of the project. These existing instruments were, however, used to help construct the Army Job Satisfaction Questionnaire (AJSQ). The AJSQ was developed to provide reliable and valid indices of satisfaction with important aspects of Army jobs and overall job satisfaction. Moreover, it was designed to be easy to understand and quick to administer.

Initial Development and Field Testing

A pool of several hundred items was assembled from previously developed military instruments related to satisfaction. Five project researchers independently reviewed the items and identified several common facets of satisfaction that might be useful to incorporate in the new instrument. A draft of the AJSQ was developed by drawing items from existing instruments to cover each of the categories of satisfaction selected for measurement. A second section of the draft AJSQ was designed to help explain differences in job satisfaction levels expressed in Section I. Items in Section II were related primarily to reasons for enlistment and reenlistment intentions, and most were drawn from other Army surveys (e.g., the Army's New Recruit Survey).

The first draft of the AJSQ was reviewed by several project researchers and then revised for field testing. Section I of the field test version comprised a total of 62 items covering six facets of satisfaction: satisfaction with supervision, co-workers, promotions, pay, work, and the Army as an organization. Section II was also revised and expanded to contain 30 items.

The field test version of the AJSQ and the short form of the Minnesota Satisfaction Questionnaire (MSQ) (Weiss, Dawis, England, & Lofquist, 1967) were administered to 271 first- and second-tour soldiers at three Army installations. The MSQ is a highly regarded 20-item measure of job satisfaction. It was used, with permission from its authors, to serve as a marker measure in this research; that is, the construct validity of the AJSQ was evaluated by examining its relationship to scores on the MSQ.

Factor analysis of the Section I satisfaction items resulted in a relatively clean representation of the expected six-factor structure. The

total number of items was reduced from 62 to 37 by eliminating items which exhibited relatively low loadings on their respective factors. Scale scores were constructed by summing the constituent items within each factor. Correlations between these scales and related scales from the MSQ are shown in Table 6.1.

The pattern of correlations demonstrates reasonable levels of discriminant and convergent validity for the AJSQ subscales. The only exception to this finding is that the MSQ scale on satisfaction with company policies and practices was more highly correlated with the AJSQ work scale than with the Army scale. The AJSQ Army scale, however, is considerably broader in concept than the MSQ Policy and Practices scale, so a particularly high correlation would not have been expected. It was simply the scale which corresponded most closely conceptually with the Army scale for satisfaction with the Army.

Table 6.1

Correlations of Army Job Satisfaction Questionnaire (AJSQ) Scales With Key Minnesota Satisfaction Questionnaire (MSQ) Scales

	AJSQ Scale					
MSQ Scale	Supervisors	Co-Workers	Promotions	Pay	Work	Army
Supervision -- Human Relations	.59	.30	.23	.10	.38	.30
Supervision -- Technical	.52	.20	.19	.36	.30	.25
Co-Workers	.10	.43	.13	.19	.21	.15
Advancement	.33	.26	.56	.25	.41	.37
Compensation	.28	.20	.27	.47	.33	.37
Variety	.29	.25	.21	.10	.52	.18
Ability Utilization	.40	.34	.33	.18	.62	.32
Achievement	.34	.34	.31	.23	.59	.30
Company Policies/ Practices	.29	.23	.30	.31	.48	.41

Note. N = 271. Correlations between similar scales are underlined.

Final Contents of the AJSQ

To conserve administration time, the AJSQ was shortened again before it was administered to soldiers in the LVI/CVII data collection. Section I of the LVI/CVII version contained 20 items related to the six facets of satisfaction with three to four items defining each facet. Section II included 13 items.

Prior to administration of the questionnaire to soldiers in LVII, the AJSQ was revised once again. In Section I, the satisfaction assessment portion, two items were added to the work subscale and a single overall satisfaction item was added as well. Whereas the other satisfaction items used a 5-point scale ranging from very dissatisfied to very satisfied, the overall item used a 7-point adjectival scale developed by Ironson and Smith (1981).

Section II was substantially revised and expanded. Four items were dropped and 25 items were added for a total of 34 items. Again, most of the items came from existing Army surveys (e.g., Army Career Satisfaction Survey, Army Families Research Questionnaire). They cover (a) reasons for enlisting, (b) reenlistment intentions, (c) unit morale, (d) perceived fairness of treatment, and (e) the impact of socio-political events (e.g., Operation Desert Storm, Army downsizing) that had occurred since the LVI AJSQ measure was administered.

A copy of the LVII version of the AJSQ is included as Appendix C. It is annotated to show the items that did not appear on the LVI/CVII version.

Psychometric Properties

AJSQ data were collected from 11,140 soldiers in the first-tour Longitudinal Validation sample (LVI), 1,025 second-tour soldiers from the Concurrent Validation sample (CVII), and 1,574 second-tour soldiers from the Longitudinal Validation sample (LVII).

As previously described, there were some differences in the two versions of the AJSQ used to collect these data; in particular, the satisfaction assessment section of the LVII form had three items that did not appear on the LVI/CVII form. The factor structure and scoring rules used for the two versions are described in this section. Descriptive statistics are provided for the two Longitudinal Validation samples (i.e., LVI and LVII). This discussion is restricted to the satisfaction items in Section I of the AJSQ.

Prior to finalizing the AJSQ scoring scheme, a principal components analysis was conducted on the 20 satisfaction items comprising the LVI/CVII version of the instrument. The results showed a very clean representation of the expected six-factor structure. Table 6.2 shows the same analysis, with similar results, conducted using the 23 items from the LVII version of the instrument. This table reflects a varimax (orthogonal) rotation; results using an oblique rotation were highly similar. Note that overall satisfaction item loads most highly (.61) on the satisfaction with work factor but also has a reasonably high loading (.40) on the satisfaction with the Army factor.

Based on the factor structure of the instrument, it seemed reasonable to construct scale scores for each of the six factors and to consider all items

Table 6.2

AJSQ LVII Principal Components Analysis: Rotated Factor Pattern Matrix

Item[a]	Satisfaction With: Supervisors	Co-Workers	Promotions	Pay	Work	Army	h^2
Item 01	**.90**	.05	.05	.03	.14	.08	.84
Item 02	**.90**	.07	.03	.03	.14	.08	.84
Item 03	**.87**	.05	.08	.05	.21	.10	.82
Item 04	.06	**.74**	.08	.04	.14	-.01	.58
Item 05	.06	**.83**	.05	.03	.15	.06	.73
Item 06	.02	**.81**	.04	.09	.15	.16	.71
Item 07	.05	.04	**.90**	.10	.10	.02	.83
Item 08	.07	.09	**.86**	.10	.13	.21	.82
Item 09	.05	.07	**.91**	.09	.13	.15	.88
Item 10	.02	.06	.11	**.89**	.09	.22	.87
Item 11	.05	.06	.10	**.87**	.04	.16	.81
Item 12	.04	.06	.08	**.89**	.08	.23	.86
Item 13	.08	.10	.02	.04	**.83**	.12	.72
Item 14	.09	.10	.05	.03	**.87**	.07	.79
Item 15	.09	.11	.10	.06	**.85**	.05	.76
Item 16	.08	.11	.09	.03	**.88**	.10	.80
Item 17	.10	.08	.10	.07	**.85**	.05	.75
Item 18	.15	.10	.07	.06	**.79**	.18	.70
Item 19	.09	.08	.19	.16	.12	**.65**	.52
Item 20	.06	.06	.04	.13	.10	**.81**	.69
Item 21	.05	.01	.03	.17	.11	**.81**	.71
Item 22	.06	.07	.14	.16	.19	**.80**	.73
Overall Satisfaction (Item 23)	.18	.13	.13	.07	.61	.40	.61
Eigenvalue	2.50	2.02	2.53	2.51	4.95	2.87	17.38

Note. h^2 = communality (sum of squared factor loadings) for variables. Loadings expected to be high on each factor are shown in bold.

[a] See Appendix C for item content.

suitable for scoring. The scores were constructed by computing the mean of the constituent items within each factor. Because the LVII version of the instrument had additional items, two sets of work subscores were computed for the LVII sample; one includes the two new items and the other does not.

An overall satisfaction composite was constructed by calculating the mean across all items appearing in Section I of the LVI/CVII version of the AJSQ. Again, two versions of the overall composite score were calculated for the LVII sample, one that includes the two additional items on the LVII form and one that does not. The LVII form also includes the single overall job satisfaction item, which was scored separately.

Table 6.3 lists all the scores that were generated from the AJSQ and their respective internal consistency reliability estimates. Table 6.4 shows the means and standard deviations for AJSQ scores. As would ordinarily be expected, satisfaction levels for second-tour soldiers tended to be higher than those for first-tour soldiers. Finally, scale intercorrelations are provided in Table 6.5. Again, as expected, scales show low to moderate correlations with each other.

Table 6.3

AJSQ Scoring System and Internal Consistency Reliability Estimates

Scale	Constituent Items[a] (LVII Numbering)	Coefficient Alpha	
		LVI	LVII
Supervision	1, 2, 3	.90	.90
Co-Workers	4, 5, 6	.80	.75
Promotions	7, 8, 9	.91	.90
Pay	10, 11, 12	.91	.90
Work	14, 15, 17, 18	.90	.90
Work II	13, 14, 15, 16, 17, 18 (LVII sample only)	--	.94
Army	19, 20, 21, 22	.84	.83
Overall Composite	Items 1-22, excluding 13 & 16	.90	.87
Overall Composite II	Items 1-22 (LVII sample only)	--	.89
Overall Satisfaction	23	--	--

[a] See Appendix C for item content.

Table 6.4

AJSQ Descriptive Statistics

Scale	LVI		LVII	
	Mean	SD	Mean	SD
Supervision	3.15	1.09	3.47	1.05
Co-Workers	3.29	.88	3.56	.78
Promotions	2.57	1.12	2.75	1.11
Pay	2.61	1.02	2.64	1.02
Work	2.79	1.06	3.27	1.09
Work II	--	--	3.30	1.07
Army	2.51	.90	2.84	.86
Overall Composite	2.80	.67	3.08	.60
Overall Composite II	--	--	3.11	.62
Overall Satisfaction[a]	--	--	4.45	1.21

Note. LVI n = 11,140; LVII n = 1,574.

[a] Note that this item is rated on a 7-point scale, whereas the other means are based on 5-point scales.

Table 6.5

AJSQ Score Intercorrelations

Scale	Supervisors	Co-Workers	Promotions	Pay	Work	Army	Comp
Supervision	--	.19	.30	.21	.38	.33	.62
Co-Workers	.16	--	.25	.17	.31	.28	.52
Promotions	.16	.18	--	.33	.38	.39	.67
Pay	.13	.17	.25	--	.30	.48	.62
Work	.32	.31	.25	.18	--	.42	.74
Army	.22	.20	.30	.42	.30	--	.75
Composite	.55	.50	.59	.57	.70	.68	--
Overall Satisfaction	.32	.28	.28	.24	.63	.47	.64

Note. Upper diagonal is LVI (n = 11,140); lower diagonal is LVII (n = 1,574).

ANALYSES, RESULTS, AND DISCUSSION

The analyses that were conducted and the results that were found are presented and discussed in turn for (a) the relationship of satisfaction to performance and (b) the prediction of turnover.

In this report, the research questions associated with satisfaction, performance, and turnover were addressed using data from the LVI sample. Related analyses using the CVII and LVII data will be described in a later report. Note also that the analyses reported in this chapter are restricted to the nine Batch A MOS because most required soldiers to have a complete array of job performance data.

Satisfaction and Performance

Sample and Measures

The sample for this set of analyses comprised 6,352 LVI soldiers from the nine Batch A MOS. In addition to the six subscores and AJSQ composite score described above, a "Challenge" subscore was computed by summing two items tapping satisfaction with job challenge and satisfaction with the opportunity to use skills and abilities.

Several indices of contextual and non-contextual performance were used in the analyses reported here. (See Chapter 1 for a complete listing of the criterion measures used in each sample and the scores that were derived from them.) The following scores came from the Army-Wide rating scales: an overall composite created by averaging across all dimensions, the overall effectiveness rating, the NCO potential rating, and three composites formed by summing across dimensions (i.e., Technical Skill and Effort, Personal Discipline, and Physical Fitness and Military Bearing). The Personal Discipline ratings composite can be considered a measure of contextual performance.

Another contextual performance composite was created by summing scores on three dimensions tapping effort, integrity, and self-development; this composite is referred to as the Contextual Army-Wide composite. A Contextual Combat Prediction Scales composite was constructed by summing scores across 10 of the 19 items on the Combat Scales. Three administrative scores from the Personnel File Form were used (i.e., Awards and Certificates, Disciplinary Actions, and M16 Qualification). Overall total scores were derived for the hands-on tests, job knowledge tests, and MOS-specific rating scales. Finally, the five factor scores from the LVI performance model were used; two of these scores (i.e., Effort/Leadership and Personal Discipline) may be considered indices of contextual performance.

Results

Table 6.6 shows sample-size weighted mean correlations between the eight job satisfaction scores and 19 job performance scores across the nine MOS. Correlations based on the measures of contextual performance (i.e., Contextual Army-wide composite, Contextual Combat Scales composite, Personal Discipline Army-wide composite, the Personal Discipline factor from the performance model, and the Effort and Leadership factor from the performance model) are printed in bold type. Both peer and supervisor performance rating data were

Table 6.6

Mean Correlations Between Job Satisfaction and Performance Weighted by Sample Size

	Army Job Satisfaction Questionnaire (AJSQ) Scales							
Performance Measure	Supervisor	Co-Workers	Promotions	Pay	Work	Army	Overall Composite	Challenge
Newly Developed Contextual Performance Variables								
Contextual Army-Wide Rating Composite	**.184**	**.090**	**.149**	**.104**	**.152**	**.144**	**.212**	**.124**
Contextual Combat Scales Composite	**.162**	**.083**	**.135**	**.092**	**.117**	**.113**	**.178**	**.091**
Army-Wide Rating Composites								
Technical Skill and Effort	.155	.088	.129	.070	.132	.115	.177	.107
Personal Discipline	**.185**	**.089**	**.156**	**.106**	**.148**	**.129**	**.209**	**.124**
Physical Fitness/Bearing	.112	.099	.143	.063	.122	.110	.165	.107
Overall Army-Wide Ratings Composite	.183	.105	.161	.093	.159	.140	.216	.131
Administrative Variables								
Awards and Certificates	.024	.030	.042	.016	.025	.035	.043	.022
Disciplinary Actions	-.083	-.048	-.117	-.090	-.107	-.098	-.140	-.099
M16 Qualification	-.013	.022	-.038	-.052	-.038	-.034	-.042	-.028
Other Measures								
Overall MOS - Specific Rating Composite	.129	.089	.108	.057	.122	.086	.152	.102
Overall Effectiveness Rating	.164	.096	.137	.073	.138	.124	.188	.115
NCO Potential Rating	.168	.103	.157	.077	.137	.127	.196	.111
Hands-on Score	.068	.023	.030	.021	.023	.041	.052	.021
Knowledge Test Score	.109	-.002	.035	.080	.018	.069	.077	.006
Performance Model Factors								
Effort and Leadership	**.148**	**.096**	**.131**	**.067**	**.126**	**.118**	**.175**	**.104**
Personal Discipline	**.174**	**.088**	**.233**	**.124**	**.159**	**.146**	**.237**	**.134**
Core Technical Proficiency	.096	.020	.039	.049	.043	.058	.077	.036
General Soldiering Proficiency	.088	-.010	.020	.066	-.005	.050	.051	-.012
Physical Fitness/Military Bearing	.071	.093	.104	.025	.085	.077	.114	.070

Note: Sample sizes range from 5,222 to 6,345. Performance rating scores are based on combined peer and supervisor rating data. Correlations based on measures of contextual performance are printed in bold type.

used in computing the correlations shown. Mean correlations across MOS are presented because analyses showed that, except for correlations based on satisfaction with promotions, most or all of the inter-occupation variance in these correlations could be attributed to sampling error. This finding is probably due to differential promotion rates across MOS.

Several points should be noted from Table 6.6. First, the correlations tend to be low.

Second, the relative magnitudes of the correlations based on different performance indices support the a priori hypotheses. The correlations between job satisfaction and contextual performance measures are higher than the correlations between job satisfaction and other performance indices. Correlations between job satisfaction and measures that are probably influenced by both contextual performance and task performance (e.g., MOS-Specific rating scale score, Disciplinary Actions, Physical Fitness/Military Bearing performance model factor score) are, in general, lower than correlations between job satisfaction and more pure measures of contextual performance, but are higher than correlations between satisfaction and "can do" measures of the technical core of the job.

Correlations based on "can do" measures of the technical core of the job (e.g., hands-on, job knowledge, and M16 Qualification scores) tend to be negligible. The difference in the correlations would most likely be even greater if the differential reliabilities of the performance measures were taken into account. The "can do" factors are not based on the rating method and are somewhat more reliable than the "will do" factors.

The only correlations that do not support the hypotheses are those based on the Awards and Certificates variable. Scores on this administrative measure should be influenced by both contextual performance and task performance and yet correlations between job satisfaction and scores on this variable are negligible.

Third, patterns of correlations vary widely by job satisfaction dimension. Examining the correlations based on the Personal Discipline performance model scores provides a good feel for the relative magnitude of correlations with various satisfaction scales. The highest correlation (.237) is based on the overall composite job satisfaction scale. In decreasing order of magnitude the other seven correlations are: promotions (.233); supervision (.174); work (.158); Army (.146); challenge (.134); pay (.124); and co-workers (.088). The relative order of the satisfaction scales in terms of the magnitudes of their correlations with other performance dimensions is similar to this for most performance dimensions.

Sample-size weighted mean correlations between job satisfaction and performance based on supervisor evaluations of performance are shown in the top half of Table 6.7. Similar correlations based on peer evaluations of performance are shown in the bottom half of the same table. A comparison of the correlations based on supervisor and peer ratings reveals consistently higher correlations between job satisfaction and supervisor evaluations of performance than between job satisfaction and peer evaluations of performance, except for the correlations based on the pay and co-workers subscales.

Table 6.7

Mean Correlations Between Job Satisfaction, Supervisor Performance Ratings, and Peer Performance Ratings Weighted by Sample Size

Performance Measure	Army Job Satisfaction Questionnaire (AJSQ) Scales: Supervisor	Co-Workers	Promotions	Pay	Work	Army	Overall Composite	Challenge
Supervisor Ratings								
Newly Developed Contextual Performance Variables								
Contextual Army-Wide Rating Composite	.192	.067	.141	.087	.145	.134	.119	.121
Contextual Combat Scales Composite	.167	.074	.127	.080	.120	.114	.175	.095
Army-Wide Rating Composites								
Technical Skill and Effort	.172	.066	.123	.061	.136	.117	.175	.112
Integrity and Self-Control	.201	.082	.147	.094	.156	.127	.209	.135
Appearance	.126	.078	.131	.062	.132	.113	.165	.118
Overall Army-Wide Rating Composite	.194	.083	.149	.080	.159	.134	.207	.136
Other Measures								
Overall MOS - Specific Rating Composite	.147	.071	.109	.055	.130	.093	.157	.111
Overall Effectiveness Rating	.176	.076	.125	.063	.138	.120	.181	.115
NCO Potential Rating	.173	.079	.152	.065	.138	.113	.186	.113
Peer Ratings								
Newly Developed Contextual Performance Variables								
Contextual Army-Wide Rating Composite	.109	.090	.116	.095	.114	.115	.162	.090
Contextual Combat Scales Composite	.098	.082	.100	.086	.082	.088	.134	.060
Army-Wide Rating Composites								
Technical Skill and Effort	.099	.098	.107	.063	.099	.097	.142	.077
Integrity and Self-Control	.122	.080	.131	.106	.107	.109	.167	.084
Appearance	.070	.108	.140	.055	.094	.090	.138	.078
Overall Army-Wide Rating Composite	.112	.106	.137	.084	.113	.112	.168	.089
Other Measures								
Overall MOS - Specific Rating Composite	.081	.095	.088	.052	.088	.070	.120	.072
Overall Effectiveness Rating	.093	.091	.108	.059	.089	.094	.134	.073
NCO Potential Rating	.108	.111	.120	.066	.098	.109	.152	.080

Note: Sample sizes range from 4,425 to 5,638.

It is possible that correlations between satisfaction and supervisor performance evaluations are higher than correlations between satisfaction and peer performance evaluations because supervisor ratings are more reliable. To explore this possibility, the correlations shown in Table 6.7 based on the three Army-wide rating composites, the overall effectiveness ratings, and the NCO potential ratings were corrected for criterion unreliability. (Interrater reliability estimates used in computing the corrected correlations ranged from .56 to .64 for the supervisor ratings and .48 to .59 for the peer ratings.) Correcting for criterion unreliability did not substantially influence the relative sizes of the correlations. Correlations based on supervisor data were still consistently higher than correlations based on peer data, with the exception of correlations based on the co-workers satisfaction scale. Corrected correlations between satisfaction with co-workers and peer evaluations of performance were higher than corrected correlations between satisfaction with co-workers and supervisor evaluations of performance.

<u>Summary</u>

The hypotheses regarding satisfaction and performance were supported by the pattern of correlations found. Sample-size weighted mean correlations between job satisfaction scores and scores on measures of contextual job performance were higher than correlations between job satisfaction and scores on other job performance measures. Sample-size weighted mean correlations between job satisfaction and measures influenced by both contextual performance and task performance were lower than correlations between job satisfaction and more pure measures of contextual performance. Correlations based on "can do" task performance measures were, in most cases, negligible.

Although the expected pattern of correlations was found, the correlations based on measures of contextual performance were lower than expected, and lower than those found in some of the previous studies of relationships between job satisfaction and OCB (e.g., Bateman & Organ, 1983; Smith et al., 1983). The very reason that led to the use of measures of contextual performance instead of measures of OCB in this study may explain the lower than expected correlations. The fact that regulations and standard procedures govern most aspects of Army life and that mechanisms exist to formally recognize adherence to those regulations and procedures makes it more likely that a person will perform prosocial acts -- in part, to avoid organizationally sanctioned punishments for failing to perform them, or to reap organizationally sanctioned benefits for performing them. Thus, performance differences resulting from differences in affective states may be more subtle in military contexts than in organizational contexts in which there are fewer strictures governing behavior.

For all satisfaction scales except the pay and co-workers scales, correlations between satisfaction and supervisor performance ratings were higher than correlations between satisfaction and peer performance ratings. These differences cannot be attributed to higher inter-rater reliabilities among the supervisor rating data than among the peer rating data. It may be that supervisors are simply more likely than peers to allow the attitudes of those they are rating to influence their performance judgments, regardless of whether these attitudes are reflected in actual performance differences. However, a more likely explanation is that supervisors are better able to discern subtle performance differences that covary with differences in satisfaction.

Prediction of Turnover

Sample and Measures

The base sample for the analyses reported here comprises 5,721 LVI soldiers representing the nine Batch A MOS. Most analyses used the overall satisfaction composite score and an overall job performance score. The overall performance score was computed by summing the five standardized factor scores.

An intention-to-reenlist variable was extracted from Section II of the AJSQ. It was scored as follows: (0) intend to leave the Army, (1) do not know, and (2) intend to reenlist. Approximately 69 percent of the soldiers in this sample (Batch A MOS; all enlistment terms) indicated that they would leave the Army, 15.5 percent indicated that they would reenlist, and 15.5 percent indicated that they were undecided.

An intention-to-attrit variable was obtained from a survey designed by the Army Family Research Program which was administered in conjunction with the LVI data collection (Research Triangle Institute, 1988). It was scored (0) if the soldier indicated that he or she was likely to leave the Army before the enlistment term expired and (1) if the soldier expected to complete his or her tour. About 4 percent of the sample reported that they intended to attrit.

It was expected that soldiers having more time remaining in their first term would be more likely to attrit, regardless of their level of job satisfaction or job performance. The fact that 3-year enlistees had a 10 percent attrition rate during the period between testing and the end of their enlistment term, compared to a 20.5 percent attrition rate for 4-year enlistees, supports this hypothesis. Base rate differences across enlistment terms were controlled by conducting turnover analyses for soldiers in each enlistment term separately.

A special variable was constructed to provide a control variable for differences in time remaining for soldiers within each enlistment term. This variable is expressed as the proportion of the soldier's term that had expired at the time that the research measures were administered. The mean proportion of term expired was .72 for 2-year term soldiers, .55 for 3-year term soldiers, and .36 for 4-year term soldiers. As expected, the mean proportion of term that had expired at the time of testing decreased as the length of the enlistment term increased.

Turnover. For most of the reenlistment analyses, reenlistees were compared to soldiers who completed their initial term of enlistment and who were eligible to reenlist. Soldiers who left the Army before completing their enlistment terms or who were otherwise ineligible to reenlist were excluded from these analyses. Excluding soldiers who were ineligible to reenlist was generally appropriate because two of the three primary predictors (intention and satisfaction) were hypothesized to be related to voluntary rather than involuntary turnover behavior. Note, however, that analyses which use job performance as the only type of predictor include soldiers who were ineligible for reenlistment. The reenlistment rate across MOS and enlistment terms, excluding those soldiers who were ineligible, was 38 percent. Adding in the soldiers who were ineligible for reenlistment, the rate dropped to 32 percent. This is substantially higher than the reenlistment intention rate which was 15

percent, a finding that is consistent with prior military research (Hiller, 1982).

For the attrition analyses, soldiers who separated prematurely for avoidable reasons were compared to those who did not. When event history analysis was used to conduct the data analysis, no observations had to be excluded from the analysis. When logistic regression analysis was used, unavoidable attrition cases (e.g., attrition due to disability) were dropped. Note that the percentage of soldiers who exited prematurely is small in this sample (13%) because attrition is most likely to occur early in the first enlistment term. The timing of the LVI data collection was such that soldiers had to have already made it well into their first term in order to be included in this analysis sample.

Bivariate correlations among the variables described above are presented in Table 6.8. As hypothesized, reenlistment intention and job satisfaction were significantly correlated with reenlistment. Job satisfaction and attrition intention were also significantly correlated with attrition. Job performance was signicantly correlated with both types of turnover-related behavior as well.

Table 6.8

Intercorrelations Among Major Turnover Analysis Variables

	JS	JP	Reenlist Intent	Attrit Intent	Reenlist	Attrit
Job Satisfaction	1.000					
Job Performance	.189	1.000				
Reenlist Intention	.268	.073	1.000			
Attrit Intention[a]	.196	.192	.099	1.000		
Reenlistment[b]	.158	.132	.369	.104	1.000	
Attrition[c]	-.089	-.263	-.060	-.264	-.263	1.000
Proportion Term Completed	-.081	.126	-*.024*	.067	*.026*	-.196

Note. n = 4,098 for all correlations involving attrition intention and 5,721 for all others. All correlations except those in italics are significant at p = .0001.

[a] Given the 96% intention-not-to-attrit base rate, the maximum possible r_{pb} is approximately .49; attrition coded 1 if soldier plans to complete term and 0 if not.

[b] Given the 32% reenlistment base rate, the maximum possible r_{pb} is about .78; reenlistment coded 1 if reenlisted and 0 if not.

[c] Given the 13% attrition base rate, the maximum possible value r_{pb} is approximately .60; attrition coded 1 if attrited and 0 if not.

Prediction of Reenlistment

<u>Satisfaction</u>. Logistic regression with stepwise selection was calculated in an effort to identify the satisfaction subscales most strongly predictive of reenlistment. However, the results were difficult to interpret because multicollinearity (moderate to high intercorrelations) among the satisfaction subscales resulted in sign reversals in some of the parameter estimates; for example, pay satisfaction appeared to be negatively related to reenlistment when in fact it is positively related. In view of the problems involved in subscale analyses, the overall job satisfaction composite was used in subsequent analyses dealing with predictions of turnover.

Performance. The relationship between job performance and reenlistment, exclusive of other predictors, was examined in two stages of analysis. First, the six indices of performance that are administrative in nature (i.e., Awards and Certificates, Disciplinary Actions, Skill Qualification Test scores, M16 Qualification, Physical Readiness test score, and Promotion Rate) were regressed onto reenlistment. These stepwise logistic regression analyses were conducted separately for each term of enlistment, and soldiers ineligible for reenlistment were included. The administrative measures were looked at separately from the for-research-only measures (i.e., hands-on tests, written job knowledge tests, and peer/supervisor ratings) because soldiers have knowledge of their level of performance on these measures which may, in turn, influence their reenlistment decision.

In the second set of analyses, the administrative measures demonstrating a significant relationship to reenlistment were combined with overall scores from the for-research-only measures to predict reenlistment. It was expected that the administrative measures would be more useful than the for-research-only measures for predicting reenlistment behavior.

With regard to the first set of analyses, none of the administrative measures met the significance criterion for entering the prediction equation in either set of analyses for the soldiers with 2-year enlistment terms. Disciplinary Actions and Promotion Rate entered the equations for both 3-year and 4-year enlistment term soldiers. In addition, the Awards and Certificates score was a significant predictor of 4-year term soldier reenlistment. The M16 Qualification score met the entry criterion for 3-year term soldiers, but it had such a weak, negative beta weight that it was not used in the second set of analyses.

The results of the second set of analyses are shown in Table 6.9. For both 3-year and 4-year term soldiers, the overall performance ratings composite (constructed using Army-wide and MOS-specific ratings dimensions) was the only for-research-only measure that entered the prediction equation. This makes sense given that supervisor evaluations help determine reenlistment eligibility. Using the correlation between the probability of reenlistment estimated by the regression equation and actual reenlistment as a gauge, the relationship between the individual performance measures and reenlistment is relatively small, with a point-biserial correlation of -.169 for 3-year term soldiers and -.183 for 4-year term soldiers.

Table 6.9

Prediction of Reenlistment Using Individual Job Performance Measures

Predictor	Parameter Estimate	Standard Error	p-value	Odds Ratio
3-Year Term (n = 3,613)				
Intercept	-4.983	.463	.0001	.007
Awards and Certificates	---	---	---	---
Disciplinary Actions[a]	.206	.051	.0001	1.229
Promotion Rate	.038	.008	.0001	1.039
Ratings	.004	.001	.0001	1.004
r_{pb} = .169[b]				
36% reenlistment[c]				
4-Year Term (n = 1,912)				
Intercept	-5.479	.718	.0001	.004
Awards and Certificates	.069	.016	.0001	1.072
Disciplinary Actions[a]	.221	.081	.0061	1.248
Promotion Rate	.034	.012	.0036	1.035
Ratings	.004	.001	.0048	1.004
r_{pb} = .183[b]				
29% reenlistment[c]				

[a] Score has been reversed-coded so relationship to reenlistment is expected to be positive.
[b] r_{pb} is the point-biserial correlation between predicted probability of reenlistment and actual reenlistment.
[c] Including soldiers ineligible for reenlistment.

Overall Analyses. The relative contribution of the three types of predictors (i.e., intention, job satisfaction, and job performance) was assessed by evaluating the improvement of model fit with the addition of each predictor to the prediction equation. Intention was the first predictor entered into the model because it is generally believed to be the most immediate precursor to turnover (e.g., Mobley et al., 1979). The overall job satisfaction composite was entered next. Overall job performance was entered last because it was not expected to be related as strongly to reenlistment as satisfaction would be. Separate reenlistment analyses were conducted for soldiers in each enlistment term (2, 3, and 4 years).

The effect of adding each variable to the model was evaluated by conducting likelihood ratio tests. A likelihood ratio test is conducted by computing the difference between the logs of the likelihood ratios for each

model (i.e., the model with the new variable and the model without the new variable). This difference, multiplied by -2, is distributed asymptotically as chi-square, with degrees of freedom in this case equal to one. In addition, an index of the predictive strength of each model was constructed by calculating the point-biserial correlation between the predicted probability of reenlistment, as determined by each model, and actual reenlistment behavior.

Results of the comparison among reenlistment prediction models are summarized in Table 6.10. For all three enlistment terms, all three models are statistically significant. However, the incremental contribution of job satisfaction and job performance over reenlistment intention appears to be minimal. There are no significant differences among the three models for 2-year enlistees. The addition of job satisfaction provides incremental predictive power over intention for 3-year enlistees, but the addition of job performance to those two predictors does not improve the model fit. In contrast, the addition of job satisfaction to the model does not significantly improve model fit for 4-year enlistees, but the addition of job performance does. Clearly, the most powerful and consistent predictor of reenlistment behavior in this sample is reenlistment intention. This conclusion is supported by examining the prediction equation parameters provided in Table 6.11.

Table 6.10

Point-Biserial Correlations Between Reenlistment and Predicted Probability of Reenlistment for Three Models

	Enlistment Term		
Alternative Models	2-Year	3-Year	4-Year
Intention	.548	.399 } b	.293
+ Satisfaction	.553 (.546)[a]	.402 (.400)	.293 (.287) } b
+ Performance	.553 (.544)	.402 (.399)	.301 (.293)
Reenlistment Rate[c]	14%	45%	39%
Maximum r_{pb}[d]	.60	.81	.80

Note. All models have chi-squares significant at $p < .05$. 2-year n = 492; 3-year n = 2,753; 4-year n = 1,322.

[a] Correlations in parentheses are adjusted for shrinkage using formula from Stein (1960).
[b] Differences in model chi-squares are significant at $p < .05$.
[c] Excluding soldiers ineligible for reenlistment.
[d] Estimated from Figure 4-5 in Nunnally (1967, p. 133).

Table 6.11

Prediction of Reenlistment Using Three Post-Enlistment Predictors

Predictor	Parameter Estimate	Standard Error	p	Odds Ratio
2-Year Term (n = 492)				
Intercept	-3.889	1.714	.0232	.020
Intention	1.990	.236	.0001	7.319
Satisfaction	.395	.266	.1375	1.485
Performance	.001	.003	.8256	1.001
14% reenlistment[a]				
3-Year Term (n = 2,753)				
Intercept	-1.494	.470	.0015	.224
Intention	1.068	.059	.0001	2.911
Satisfaction	.157	.066	.0169	1.170
Performance	.000	.000	.5479	1.001
45% reenlistment[a]				
4-Year Term (n = 1,322)				
Intercept	-2.571	.681	.0002	.076
Intention	.779	.080	.0001	2.179
Satisfaction	.015	.094	.8730	1.015
Performance	.003	.001	.0118	.084
39% reenlistment[a]				

[a] Excluding soldiers ineligible for reenlistment.

The point-biserial correlations between the predicted probability of reenlistment and actual reenlistment suggest that the reenlistment behavior of 3-year enlistees is predicted more successfully than that of the 4-year enlistees. Despite a relatively low reenlistment base rate, the reenlistment behavior of 2-year enlistees is most accurately predicted. This finding is particularly interesting given that the maximum point-biserial correlation possible with the 2-year term's 14 percent reenlistment base rate (r_{pb}=.60) is considerably lower than that possible with the 3-year term's 45 percent base rate (r_{pb}=.81). It is reasonable to expect, however, that the more distant the soldier is from making a reenlistment decision at the time of testing, the less accurate the prediction is going to be. Moreover, some 2-year term enlistees may have already made their reenlistment decisions prior to the time of testing.

Prediction of Attrition

Two sets of analyses were used to model attrition. The first set used logistic regression analysis to assess the incremental contribution of four variables to the prediction of avoidable, late-term attrition. These variables were (a) the proportion of term completed when measures were administered, (b) intention to attrit, (c) overall job satisfaction, and (d) overall job performance.

In the second set of analyses, event history analysis was used to determine whether job satisfaction contributed significantly to model fit over the seven best predictors identified by the analyses reported in Chapter 5 of this report (i.e., high school diploma status, ASVAB Quantitative composite, and the following scores from the ABLE: Nondeliquency, Dominance, Physical Condition, Self-Esteem, and Social Desirability).

In both sets of analyses, the event of interest was avoidable attrition (both voluntary and involuntary) as described in Knapp (1993). All other turnover-related outcomes (i.e., reenlistment, completion of first term, unavoidable attrition) were treated as censored observations in the event history analyses. In the logistic regression analyses, unavoidable attritions were dropped from the analysis sample.

Analyses were performed separately for soldiers with 3-year and 4-year enlistment terms. Due to small sample sizes, data from soldiers with 2-year terms of enlistment were not analyzed. As expected, the base rate of attrition was markedly higher for 4-year enlistees (20.5%) than for 3-year enlistees (10%). For the event history analyses, analyses were also conducted separately for combat and non-combat MOS and were stratified by MOS. This is the same analysis approach described in Chapter 5 except that only soldiers in Batch A MOS were included.

Intention, Satisfaction, and Performance. The first set of analyses were intended to assess the incremental contribution of four variables to the prediction of avoidable attrition. This was done by comparing the overall fit of the following four nested models: (a) proportion of term completed when measures were administered, (b) proportion of term completed and intention to attrit, (c) proportion of term completed, intention, and satisfaction, and (d) proportion of term completed, intention, satisfaction, and performance. The proportion of term completed when the intention, satisfaction, and performance measures were administered was considered a statistical control variable and was therefore always entered first into the model. The other variables were entered into the model in the order of the expected strength of their relationship to attrition (i.e., intention, satisfaction, performance).

The four models are compared in Table 6.12. All models were statistically significant ($p < .05$) for both 3- and 4-year enlistees. Likelihood ratio tests of the model chi-squares indicate that the addition of each predictor provided incremental predictive power for 4-year term enlistees, and for all predictors except job satisfaction for the 3-year term enlistees. There was little difference between enlistment terms in the predictive power of the full model. The correlations between the predicted probability of attrition and attrition were -.372 (3-year term) and -.370

Table 6.12

Point-Biserial Correlations Between Attrition and Predicted Probability of Attrition for Four Models[a]

	Enlistment Term						
Alternative Models	3-Year			4-Year			
Proportion of term completed	-.115		[b]	-.144		[b]	
+ Intention	-.300	(-.297)		-.258	(-.250)		[b]
+ Satisfaction	-.302	(-.298)	[b]	-.281	(-.271)	[b]	
+ Performance	-.370	(-.366)		-.372	(-.363)		
Avoidable attrition rate	10%			20.5%			
Maximum r_{pb}[c]	.55			.70			

Note. All models have chi-squares significant at $p < .05$. 3-year n = 2,540; 4-year n = 1,197.

[a] Correlations in parentheses are adjusted for shrinkage using formula from Stein (1960).
[b] Differences in model chi-squares are significant at $p < .05$.
[c] Estimated from Figure 4-5 in Nunnally (1967, p. 133).

(4-year term). Parameter estimates for the model including all four predictors are provided in Table 6.13.

Pre-Enlistment Predictors and Satisfaction. Results of the event history analyses examining pre-enlistment predictors and job satisfaction as predictors of attrition are provided in Tables 6.14 through 6.17. Whether job satisfaction provides incremental model fit over the pre-enlistment predictors was determined using likelihood ratio tests. In all cases (3-year term, combat MOS; 4-year term, combat MOS; 3-year term, non-combat MOS; and 4-year term, non-combat MOS), satisfaction significantly improved model fit.

Examination of the equation parameters reveals that, unlike the analyses presented in Chapter 5, most of the pre-enlistment predictors were not significantly or consistently related to attrition. Whereas all seven of these predictors were related to attrition in earlier analyses, the ABLE Nondelinquency composite and job satisfaction were the only predictors consistently shown to be significantly related to attrition in these analyses.

Table 6.13

Prediction of Attrition Using Post-Enlistment Predictors

Predictor	Parameter Estimate	Standard Error	*p*	Odds Ratio
3-Year Term (n = 2,540)				
Intercept	-8.220	.760	.0001	.000
Prop. term completed	2.796	.527	.0001	16.375
Intention	2.022	.261	.0001	7.557
Satisfaction	.003	.109	.9765	1.003
Performance	.015	.001	.0001	1.015
10% avoidable attrition				
4-Year Term (n = 1,197)				
Intercept	-7.888	.842	.0001	.000
Prop. term completed	3.021	.820	.0002	2.509
Intention	1.251	.281	.0001	3.494
Satisfaction	.278	.120	.0202	1.320
Performance	.013	.002	.0001	1.013
20.5% avoidable attrition				

Table 6.14

Prediction of Attrition Using Pre-Enlistment Predictors: Combat, 3-Year Job Group (C3)[a]

Predictor	Parameter Estimate	Standard Error	p	Odds Ratio
HSDG	-.471	.202	.0197	0.624
QUANTITATIVE Composite	-.010	.006	.1252	0.990
Physical Condition Scale	.010	.031	.7577	1.010
Dominance Scale	.018	.023	.4239	1.019
Self-Esteem Scale	-.012	.029	.6757	0.988
Nondelinquency Scale	-.037	.016	.0252	0.964
Social Desirability Scale	.019	.025	.4457	1.020
Satisfaction	-.410	.115	.0004	0.663

	-2 Log L	χ^2 (df=1)
Pre-enlistment Predictors	1946.17	
+ Satisfaction	1933.57	12.6*

* Significant at $p < .01$.

[a] n = 1,609. Overall avoidable attrition rate is 11%.

Table 6.15

Prediction of Attrition Using Pre-Enlistment Predictors: Combat, 4-Year Job Group (C4)[a]

Predictor	Parameter Estimate	Standard Error	p	Odds Ratio
HSDG	-.199	.319	.5329	0.819
QUANTITATIVE Composite	-.004	.009	.5981	0.995
Physical Condition Scale	.079	.043	.0630	1.083
Dominance Scale	.077	.031	.0142	1.080
Self-Esteem Scale	-.023	.038	.5489	0.977
Nondelinquency Scale	-.109	.022	.0001	0.896
Social Desirability Scale	.032	.038	.3966	1.032
Satisfaction	-.563	.173	.0011	0.570

	-2 Log L	χ^2 (df=1)
Pre-enlistment Predictors	892.62	
+ Satisfaction	881.72	10.9*

* Significant at $p < .01$.

[a] n = 526. Overall avoidable attrition rate is 17%.

Table 6.16

Prediction of Attrition Using Pre-Enlistment Predictors: Non-Combat, 3-Year Job Group (NC3)[a]

Predictor	Parameter Estimate	Standard Error	p	Odds Ratio
HSDG	-.567	.208	.0065	0.567
QUANTITATIVE Composite	-.020	.006	.0008	0.980
Physical Condition Scale	-.002	.026	.9469	0.998
Dominance Scale	.044	.021	.0372	1.045
Self-Esteem Scale	-.043	.025	.0918	0.958
Nondelinquency Scale	-.032	.015	.0313	0.969
Social Desirability Scale	.050	.023	.0281	1.051
Satisfaction	-.460	.106	.0001	0.631

	-2 Log L	χ^2 (df=1)
Pre-enlistment Predictors	2204.17	
+ Satisfaction	2185.30	18.87*

* Significant at $p < .01$.

[a] n = 1,836. Overall avoidable attrition rate is 11%.

Table 6.17

Prediction of Attrition Using Pre-Enlistment Predictors: Non-Combat, 4-Year Job Group (NC4)[a]

Predictor	Parameter Estimate	Standard Error	p	Odds Ratio
HSDG	-.314	.254	.2168	0.731
QUANTITATIVE Composite	-.006	.007	.4320	0.994
Physical Condition Scale	-.049	.031	.1088	0.952
Dominance Scale	.021	.022	.3570	1.021
Self-Esteem Scale	.003	.030	.9247	1.003
Nondelinquency Scale	-.044	.017	.0098	0.956
Social Desirability Scale	.037	.028	.1870	1.038
Satisfaction	-.325	.125	.0093	0.723

	-2 Log L	χ^2 (df=1)
Pre-enlistment Predictors	1542.07	
+ Satisfaction	1535.40	6.67*

* Significant at $p < .01$.

[a] n = 674. Overall avoidable attrition rate is 22%.

The difference between the two sets of analyses is that those reported in Chapter 5 pertain to attrition occurring throughout the first enlistment term, and the analyses reported here relate only to attrition that occurs relatively late in the enlistment term. These findings support the notion that late-term attrition and the more common early-term attrition are not equivalent criteria. They also show that, despite this apparent nonequivalence, the Nondelinquency measure retains substantial predictive power throughout the first term of enlistment.

SUMMARY

The AJSQ is a reliable and valid measure of satisfaction and it has added another dimension to the Project A/Career Force research program. Satisfaction is related to contextual measures of performance, though not as strongly as anticipated. (Both satisfaction and performance are related to reenlistment, but much of their predictive power is eroded when reenlistment intention is added into the prediction equation.) The findings regarding attrition are similar to those associated with reenlistment, although performance retains appreciable predictive power even when intention to attrit is included. Satisfaction does predict attrition very well compared to pre-enlistment cognitive ability and temperament predictors.

A final observation is that the length of the enlistment term and the point in that term at which attitudes and performance are measured have a substantial impact on the ability to predict turnover-related behavior.

This chapter has focused on relationships between job satisfaction, job performance, and two types of military turnover (reenlistment and attrition) in the first tour. Using Project A/Career Force data, an extensive analysis of job satisfaction as a criterion measure to be predicted has been reported by Carter (1991). Additional analyses in the future will be concerned with the consistency of individual differences in job satisfaction over time and with the influence of different frames of reference on individual job satisfaction.

Chapter 7

OVERALL SUMMARY AND FUTURE PLANS

John P. Campbell

SUMMARY OF YEAR FOUR

During the fourth year of the Career Force Project, two major sets of analyses related to predicting second-tour performance were carried out. The first was the basic validation of the ASVAB and the Experimental Battery against the Longitudinal Sample second-tour performance measures (LVII). Much larger sample sizes were available for LVII than for the earlier analyses of the CVII sample and the predictor information was much more complete. The second set of LVII analyses examined how accurately second-tour performance could be predicted from first-tour performance and also from performance during training. The similarity in the latent structure of performance at each of the three levels (i.e., training, first tour, second tour) permits the individual performance factors to be examined in terms of their divergent and convergent relationship across levels.

In addition to the analyses of the LVII data, the LVI data base was used for three major purposes. First, an extensive analysis of the absolute and discriminant validity for optimal combinations of the ASVAB subtests plus the Project A ECAT battery was carried out, using software that compared all possible combinations of tests at each stage. Second, the full Longitudinal Validation predictor sample was used to examine the relationship of the Experimental Predictor Battery to first-term attrition, using event history (survival) analysis. Finally, the extensive Project A/Career Force data base generated by the Army Job Satisfaction Questionnaire (AJSQ) was used to estimate the relationship of job satisfaction to the components of first-tour performance, second-tour performance, first-term attrition, and reenlistment/non-reenlistment.

LVII Basic Validation

A complete LVII validation required that a reasonable subset of the soldiers in the original LV predictor sample (N = 28,000 for the nine Batch A MOS) finish their first tour of duty successfully, reenlist for a second tour, and be available at a designated set of data collection sites for performance assessment during the spring and summer of 1992 (i.e., approximately the third year of their second tour). Sufficient data on both the predictor battery and the criterion measures were available for seven of the nine Batch A MOS, and the total sample size was approximately 1,200.

Multiple correlations were computed for each predictor set (ASVAB, spatial, computer, ABLE, AVOICE, JOB) against each of 13 criterion scores (the six LVII performance factors, two modifications of the Leadership factor score, a "will do" composite, a "can do" composite, the total of the six factor scores, and separate scores for the Hands-On and Job Knowledge tests).

In general, the validity of the ASVAB is as high for predicting second-tour performance as it is for predicting first-tour performance. Its highest validities are for predicting Core Technical Performance, but it also has high validity for predicting second-tour leadership performance (average R = .50),

even when the Leadership criterion score is purged of all possible method variance.

The Personal Discipline and the Fitness and Bearing factors are not predicted as well in second tour as they were in first tour by any of the tests in the predictor battery. Possible explanations are the restriction in range on these criterion variables (unrelated to ASVAB) for the more highly selected second-tour samples and the change in the meaning of the indices for more highly selected and experienced soldiers.

The validity of the ABLE (administered at the time of accession) for predicting the Leadership factor in LVII was about the same as the ABLE validity for predicting the more heterogeneous ELS factor in CVI. Finally, the incremental validity of the Experimental Battery over ASVAB was not as great in the LVII sample as it was in CVI and LVI.

Prediction of Future Performance From Past Performance

The convergent and divergent validity of specific components of current performance for predicting future performance was examined by computing three intercorrelation matrices. Respectively, the three matrices were (a) the intercorrelations of training performance criterion measures with first-tour performance measures, (b) the intercorrelations of first-tour performance measures with second-tour performance measures, and (c) the intercorrelations of training performance measures with second-tour performance measures (i.e., six years later). The variables in each matrix included the simple sum factor scores from each performance model, the "can do," "will do," and total factor composite scores, and a number of scoring variations that attempted to equate the content of the factor scores precisely across cohorts. Also of interest was the validity with which early-stage peer and supervisory assessment of NCO potential can predict subsequent NCO performance during the second tour.

The results exhibited a striking degree of convergent and divergent validity across the three organizational levels. In addition, the magnitude of the correlations of performance with performance was quite high. Performance in training does predict performance on the job (first tour), and entry-level performance does predict second-tour NCO performance, and with considerable differential validity across performance components. There are also substantial correlations between performance in training (AIT) and performance as an NCO six years later.

Finally, a series of stepwise analyses of the predictor battery versus current performance as predictors of future performance were carried out. The results clearly showed that the test battery and measures of current performance each contribute considerable unique variance to the prediction of future performance.

In general, the cross-level analyses of the Project A/Career Force performance measures provide considerable additional evidence for their construct validity.

ECAT Optimal Battery Analysis

The Enhanced Computer Administered Test Battery (ECAT) analyses were conducted using Project A/Career Force data, to assist the Manpower Accession

Policy Working Group in its deliberations about possible revisions to the ASVAB. The overall purpose was to examine seven tests from the Project A/Career Force Experimental Battery in terms of their potential usefulness as additions to, or replacements for, the ten subtests of the current ASVAB. The seven candidate tests were:

1) Spatial: Assembling Objects
2) Spatial: Orientation
3) Spatial: Reasoning
4) Short-Term Memory (computerized)
5) Target Identification (computerized)
6) One-Hand Tracking (computerized)
7) Two-Hand Tracking (computerized)

The general procedure was to first specify that the Arithmetic Reasoning and Word Knowledge subtests of the ASVAB would be the base components of every potential battery. Then all possible combinations of a specific number of the remaining 15 tests (8 ASVAB and 7 ECAT) would be evaluated on several indices. For example, if 9 tests were to be added, all possible combinations of 9 tests from the pool of 15 were evaluated. The number of additional tests to be considered was determined by the desired battery length (in terms of battery administration time). Three different battery lengths were evaluated (74-104 minutes, 134-164 minutes, 194-224 minutes); the longest interval corresponds to the current ASVAB administration time.

The indices computed for each unique test battery were as listed below. All of the indices are based on the data generated by computing multiple correlations of the battery with a particular criterion in each of the Batch A MOS. That is, the regression weights are computed within MOS each time. For each battery there are nine such estimates. Each set of MOS specific weights for the battery was then applied to the data in each of the other Batch A MOS to compute a generalized validity estimate. There are 9 x 8, or 72, such estimates. The indices are defined as follows:

- Mean absolute validity is the average of the nine MOS specific estimates after they have been corrected by the Rozeboom formula.

- Mean generalizable validity is the average of the 72 off-diagonal values obtained when weights computed on one MOS are applied to the data from the other eight MOS to compute a correlation between the criteria and a weighted predictor composite.

- Discriminant validity is the difference between mean absolute validity and mean generalizable validity for a specific test battery.

- The Brogden Index of Classification Efficiency (BCE) is computed as $R\sqrt{1 - \bar{r}}$, when R is the mean absolute validity across jobs and r is the mean intercorrelation among the predicted criterion scores for each job. If the selection ratio and the number of jobs remain constant, then BCE is isomorphic with, but not in the same metric as the gain in mean predicted performance (MPP) that results from classification.

- Subgroup difference is the difference in mean predictor composite scores for Whites/Blacks, Whites/Hispanics, and Men/Women, computed on the combined sample for each of the nine sets of MOS-specific weights and then averaged over the nine estimates of the differences.

The entire set of analyses was carried out for each of three criteria: the TECH score from the End-of-Training (EOT) School Knowledge test; the Core Technical Proficiency (CTP) factor score for first-tour performance; and the Hands-On test score. The discriminant validity index and the Brogden Classification Efficiency Index were clearly related.

There was a positive correlation among all the battery indices. That is, the greater the absolute validity of a battery, the greater the discriminant validity and the greater the subgroup differences. In general, the White/Black differences were more highly correlated with absolute validity than with discriminant validity, and also classification efficiency, but the reverse was true for Male/Female differences.

The effects of increasing battery length were relatively small. Absolute validity increased the most between the short and medium length batteries while discriminant validity increased slightly but uniformly as a function of increasing test length.

The ASVAB subtests made their greatest contribution to absolute validity rather than discriminant validity and the ECAT tests tended to make their greatest contribution to discriminant validity. When maximizing discriminant validity, the ASVAB subtests that were replaced most often by ECAT tests were General Science, Mechanical Knowledge, and Numerical Operations. The ECAT tests most often substituted for ASVAB subtests were Assembling Objects, Target Identification, and One-Hand Tracking.

The most general conclusion that can be drawn from the data is that all the relevant indices cannot be maximized by the same test battery. Although there is a relatively high correlation between absolute validity and discriminant validity, it is not unity and maximizing discriminant validity will detract somewhat from absolute validity. Choosing a battery to reduce subgroup differences will also reduce validity. Consequently, in the context of any decision about how to reconstitute a test battery, a number of very important trade-offs must be made.

Modeling First-Term Attrition

Attrition, or turnover, from an organization is a time-dependent outcome. That is, the relative frequency of attrition varies over time in ways that appear to be lawful and explainable, rather than unsystematic or random. These time-dependent frequencies invite questions about the differential prediction of attrition across time and whether the antecedents of attrition are individual differences, situational properties, and/or their interaction.

The individual differences component of these time-dependent relationships was examined by using event history analysis to model the relationship of the ASVAB and the Experimental Battery predictors with first-tour attrition for the Longitudinal Validation Batch A sample (i.e.,

approximately 28,000 soldiers from nine MOS in the 1986/87 accession cohort). Event history models look at the degree to which information about individual or situational differences can account for the survivor function (the proportion of the initial sample that remains in the organization as a function of the passage of time) and the hazard function (the relative rate at which people are leaving the organization as a function of time).

Given a baseline survivor function, the variable that currently and historically exerts the greatest effect on attrition rates is high school graduation status. The incremental predictive effects of the ASVAB and each predictor type in the Experimental Battery were analyzed for four subgroups of soldiers defined by combat vs. non-combat MOS and 3-year vs. 4-year enlistments. The ASVAB, ABLE, and AVOICE each contributed to predicting the attrition over and above the effects of high school graduation status.

A multiple stepwise regression analysis was used to identify the best fitting predictor composite. It consists of high school graduation status, five ABLE scales (Nondelinquency, Dominance, Physical Condition, Self-Esteem, and Social Desirability), and the Quantitative composite from ASVAB. Using progressively higher cut scores on this composite, it was possible to demonstrate very substantial changes in the survivor function for all four subgroups of the LV sample. Perhaps the most interesting interaction among the predictors was that the ABLE contributes significantly more predictive validity for high school graduates than for non-high school graduates, which is counter to the conventional wisdom on this issue.

In general, the information in the Experimental Battery provides significant incremental validity to high school graduation status as a predictor of survival rates in the first tour. It remains to be determined whether the available predictors can account for specific differences in the attrition rate across time. For example, do different things account for attrition during training as contrasted to attrition during the second and third years of a 4-year term of enlistment?

Job Satisfaction as a Predictor of Attrition and Reenlistment

The Army Job Satisfaction Questionnaire (AJSQ) uses 22 items in a Likert-type response format to assess six facets of individual job satisfaction. Thirty-four additional items reflect the individual's background and frame of reference. The AJSQ was administered to all individuals in the LVI and LVII samples. For both samples, scores on the AJSQ were related to the available performance measures and to the soldiers' stated intentions to attrit or to reenlist.

In the analysis of the LVI sample, the job satisfaction scores were related to actual attrition and actual reenlistment. In this sample the AJSQ was administered 1 to 2 years before the individual's first term was to be concluded. Consequently, most of the cohorts' attrition had already taken place and the subsequent attrition rate was relatively low.

In general, the correlations of job satisfaction with job performance and reenlistment were significant but low. For performance, the Rs were about .20, which is consistent with previous Army research. The job satisfaction/ job performance relationship in the military Services is consistently in the .20 - .25 range while the analogous results in the non-military sector are in

the .10 - .15 range. The correlations with attrition were lower, but the low base rate for attrition tempers any conclusions that might be drawn. The relationship of job satisfaction to the stated intent to reenlist is much higher, and intention is related to actual reenlistment to a significant degree.

FUTURE PLANS

All major data collections that were designed as part of Project A and the Career Force Project have been completed; all major data files are edited and in place; and, with the completion of the current report, all "basic" predictor and criterion analyses have also been completed. What remains are a series of maximizing/optimizing analyses as regards selection efficiency and classification efficiency, and the development of a procedure for estimating potential versus actual gains from classification under a variety of psychometric, organizational, and labor market conditions. The planned sequence of events is as follows.

The ASVAB plus the entire Experimental Battery will be used to identify the 45 equations that maximize predictive validity (i.e., selection validity) for the five LVI performance factors in each of the nine Batch A MOS. These 45 equations will then be examined in terms of their differential validity across MOS and across performance components within MOS. One major question to be addressed concerns the extent to which these 45 predictor equations can be collapsed across performance components within MOS, and across MOS within performance components, without losing predictive information. Once that question is answered, the next issue concerns the loss of predictive information that is incurred when validity weights, synthetic validity weights, and unit weights are substituted in turn for the sample-based regression weights.

To the fullest extent possible, the above examination of the content and properties of the optimal selection battery for each major performance component in each MOS will also be carried out for predicting second-tour performance. However, the LVII sample is much smaller than LVI, which places more limitations on the analyses.

The optimal equations identified in the above analyses will also be evaluated in terms of their selection "fairness." The Career Force Project has underway a major analysis of prediction fairness that uses the regression model to examine virtually all predictor/criterion combinations in each MOS for their degree of under- and over-prediction for minority groups. The report of these analyses will be a stand-alone document that thoroughly portrays the extent of subgroup differences, predictive fairness, and predictive validity in all relevant situations.

The optimal equation analyses will also be used to guide the analysis of the full "Roll-Up Model" which was first described in the Statement of Work (SOW) for Project A. That is, what is the effect on the prediction of future performance as an NCO as more and more information becomes available on an individual? Consequently, the validities for predicting each second-tour performance component will be analyzed for (a) ASVAB alone; (b) ASVAB plus the Experimental Battery (EB); (c) ASVAB + EB + end of training performance (EOT); (d) ASVAB + EB + EOT + first-tour performance (LVI). Each set of information

will be analyzed in terms of its unique and cumulative contribution to the prediction of NCO performance.

Still in the context of estimating the maximum potential selection validity that the predictive battery provides, two additional sets of analyses are being conducted. One is directed at estimating the effects on validity of using empirical keys for the ABLE and AVOICE. Each inventory is being analyzed against the Core Technical Performance factor, the Leadership factor, and attrition with all MOS for which sample sizes are sufficient. Three different types of keys are being compared and they will be evaluated in terms of cross-validation estimates, cross-MOS generalization/differentiation, and cross-cohort (i.e., CVI vs. LVI) stability.

The second set of analyses concerns correcting all major Project A/Career Force validity estimates for criterion unreliability. In the final analysis we would like to make comparative statements about the population validities for predicting different performance components, or validities for the same performance component when it is measured in different ways. Doing this requires removing differences in estimates that are due simply to differences in criterion reliabilities.

The culmination of the Career Force Project will be a summarization of its implications for the field of selection and classification research in general, and for the Army's selection and classification system in particular. In this regard, the Project has underway a major effort to develop a comprehensive procedure for estimating the maximum potential gain from classification under varying conditions, and for evaluating the extent to which alternative operational personnel assignment systems can realize the maximum potential gain, and at what cost.

We will attempt to integrate previous work on classification efficiency estimation, personnel assignment systems, utility measurement, and cost/performance trade-off models into an overall design for an evaluation test bed that uses the extensive Project A/Career Force data base as the specifications for its latent structure. If successful, the design for such a test bed would permit systematic evaluations of changes in the test battery, changes in costs, changes in job assignment procedures, and so forth in terms of a realistic simulation of their effects on meaningful classification outcomes.

Along the way, a number of critical research questions must be investigated. For example, what is the most reasonable way to select predictors for a classification battery? Or, what is the most appropriate index with which to represent classification gain? Such an effort is a difficult undertaking. However, it represents a high payoff from a very comprehensive series of projects and is extremely important, for both scientific and practical purposes.

References

Abrahams, N. M., Pass, J. J., Kusulas, J. W., Cole, D. R., & Kieckhaefer, W. F. (1993). Incremental validity of experimental computerized tests for predicting training criteria in military technical schools (Contract N66001-90-D-9502, Delivery Order 7J13). San Diego: Navy Personnel Research and Development Center.

Ackerman, P. L. (1989). Within task intercorrelations of skilled performance: Implications for predicting individual differences? (A comment on Henry & Hulin, 1987). Journal of Applied Psychology, 74, 360-364.

Allison, P. D. (1984). Event history analysis: Regression for longitudinal event data (Sage University paper series on quantitative applications in the social sciences, No. 07-046). Beverly Hills, CA: Sage.

Austin, J. T., Humphreys, L. G., Hulin, C. L. (1989). Another view of dynamic criteria: A critical reanalysis of Barrett, Caldwell, and Alexander. Personnel Psychology, 42, 583-596.

Barrett, G. V., & Alexander, R. A. (1989). Rejoinder to Austin, Humphreys, and Hulin: Critical reanalysis of Barrett, Caldwell, and Alexander. Personnel Psychology, 42, 597-612.

Barrett, G. V., Alexander, R. A., & Doverspike, D. (1992). The implications for personnel selection of apparent declines in predictive validities over time: A critique of Hulin, Henry, and Noon. Personnel Psychology, 45, 601-617.

Barrett, G. V., Caldwell, M. S., & Alexander, R. A. (1985). The concept of dynamic criteria: A critical reanalysis. Personnel Psychology, 38, 41-56.

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship." Academy of Management Journal, 26, 587-595.

Boesel, D., & Johnson, K. (1984). Why Service members leave the military: Review O. of the literature and analysis. Washington, DC: Defense Manpower Data Center.

Borman, W.C., & Bleda, P.R. (1978). Measuring motivation and job satisfaction in a military context (ARI TP 309). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt and W. C. Borman (Eds.), Personnel selection in organizations. San Francisco: Jossey-Bass.

Brogden, H. (1959). Efficiency of classification as a function of number of jobs, percent rejected, and the validity and intercorrelation of job performance estimates. Educational and Psychological Measurement, 19(2), 181-190.

Campbell, C. H., Ford, P., Rumsey, M. G., Pulakos, E. D., Borman, W. C., Felker, D. B., de Vera, M. V., & Riegelhaupt, B. J. (1990). Development of multiple job performance measures in a representative sample of jobs. Personnel Psychology, 43, 277-300.

Campbell, J. P. (Ed.) (1987). Improving the selection, classification, and utilization of Army enlisted personnel: Annual report, 1986 fiscal year (ARI Technical Report 813101). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Campbell, J. P. (Ed.) (1989). Improving the selection, classification, and utilization of Army enlisted personnel: Annual report, 1987 fiscal year (ARI Technical Report 862). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 219 046)

Campbell, J. P. (Ed.) (1991). Improving the selection, classification, and utilization of Army enlisted personnel: Annual report, 1988 fiscal year (ARI Research Note 91-34). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 233 750)

Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modeling job performance in a population of jobs. Personnel Psychology, 43, 313-333.

Campbell, J. P., & Oppler, S. (1990). Modeling of second-tour performance. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 fiscal year (ARI Technical Report 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 252 675)

Campbell, J. P., & Zook, L. M. (Eds.) (1990). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 fiscal year (ARI Technical Report 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 252 675)

Campbell, J. P., & Zook, L. M. (Eds.) (1991). Improving the selection, classification, and utilization of Army enlisted personnel: Final Report on Project A (ARI Research Report 1597). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 242 921)

Campbell, J. P., & Zook, L. M. (Eds.) (1994a). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1991 fiscal year (ARI Research Note 94-10). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 278 726)

Campbell, J. P., & Zook, L. M. (Eds.) (1994b). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1992 fiscal year (ARI Research Note 94-27). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B188 259)

Campbell, R. V. (1985). Scorer training materials (ARI RS-WP-85). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Carter, G. W. (1991). A study of relationships between measures of individual differences and job satisfaction among U.S. Army enlisted personnel. Unpublished doctoral dissertation, University of Minnesota.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid diagnostic signs. Journal of Abnormal Psychology, 74, 271-280.

Claugy, J. G. (1978). Multiple regression and validity estimation in one sample. Applied Psychological Measurement, 2, 295-801.

Collett, D. (1991). Modelling binary data. London: Chapman & Hall.

Cox, D. R. (1972). Regression models and life tables. Journal of the Royal Statistical Society, 34, 187-202.

Darlington, R. B. (1972). Multiple regression in psychological research and practice. Psychological Bulletin, 69(3), 161-182.

Department of Defense (1982). Profile of American youth: 1980 nationwide administration of the Armed Services Vocational Aptitude Battery. Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics).

Department of Defense (1985, May). Defense manpower quality (Volumes I-III). Report to the House and Senate Committees on Armed Services. Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Installations, and Logistics).

Equal Opportunity Employment Commission, U.S. Civil Service Commission, Department of Labor, & Department of Justice (1978, August). Uniform guidelines on employee selection procedures. 43 Fed. Reg. 166, 38290-38309.

Etheridge, R. M. (1989). Family factors affecting retention: A review of the literature (ARI Research Report 1511). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Farkas, A. J., & Tetrick, L. E. (1989). A three-wave longitudinal analysis of the causal ordering of satisfaction and commitment on turnover decisions. Journal of Applied Psychology, 74, 855-868.

Fichman, M. (1988). Motivational consequences of absence and attendance: Proportional hazard estimation of a dynamic motivation model. Journal of Applied Psychology, 73, 119-134.

Fichman, M. (1989). Attendance makes the heart grow fonder: A hazard rate approach to modeling attendance. Journal of Applied Psychology, 74, 325-335.

Flyer, E. S. (1959). Factors relating to discharge for unsuitability among 1956 airmen accessions to the Air Force (WADC-TN-59-201). Lackland Air Force Base, TX: Personnel Research Laboratory.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), Review of personality and social psychology (Vol. 2, pp. 141-165). Beverly Hills, CA: Sage.

Hanson, M. A., & Borman, W. B. (in press). Development and construct validation of the Situational Judgment Test (SJT) (ARI Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hanson, M. A., Campbell, J. P., & McKee, A. S. (1994). Development of the longitudinal validation sample second-tour performance model. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1992 fiscal year (ARI Research Note 94-27). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B188 259)

Harrell, F. (1980). The PHGLM procedure. In SAS Supplemental Library User's Guide (pp. 119-131). Cary, NC: SAS Institute.

Harrison, D. A., & Hulin, C. L. (1989). Investigations of absenteeism: Using event history models to study the absence-taking process. Journal of Applied Psychology, 74, 300-316.

Henry, R. A., & Hulin, C. L. (1987). Stability of skilled performance across time: Some generalizations and limitations on utilities. Journal of Applied Psychology, 72, 457-462.

Hiller, J.R. (1982). Analysis of second term reenlistment behavior. Santa Monica, CA: Rand Corporation.

Hom, P. W., & Hulin, C. L. (1981). A competitive test of the prediction of reenlistment by several models. Journal of Applied Psychology, 66, 23-39.

Hom, P. W., Katerberg, R., & Hulin, C. L. (1979). Comparative examination of three approaches to the prediction of turnover. Journal of Applied Psychology, 64, 280-290.

Hom, P. W., Prussia, G. E., & Griffeth, R. W. (1992). A meta-analytical structural equations analysis of a model of employee turnover. Journal of Applied Psychology, 77, 890-909.

Hough, L. M. (1992). The "Big Five" personality variables - construct confusion: Description versus prediction. Human Performance, 5, 139-155.

Hough, L. M., Barge, B. N., & Kamp, J. D. (1987). Non-cognitive measures: Pilot testing. In N. G. Peterson (Ed.), Development and field test of the trial battery for Project A (ARI Technical Report 739). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. Journal of Applied Psychology, 75(5), 581-595.

Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. Psychological Bulletin, 107, 328-340.

Iaffaldano, M. T., & Muchinsky, P. M. (1985). Job satisfaction and job performance: A meta-analysis. Psychological Bulletin, 97, 251-273.

Ironson, G. H., & Smith, P. C. (1981). Anchors away - The stability of meaning when their location is changed. Personnel Psychology, 34, 249-262.

Jensen, A. R. (1985). The nature of black-white differences on various psychometric tests. The Behavioral and Brain Sciences, 29, 297-300.

Johnson, C. D., & Zeidner, J. (1990). Classification utility: Measuring and improving benefits in matching personnel to jobs (IDA Paper P-2240). Alexandria, VA: Institute for Defense Analysis.

Jöreskog, K. G., & Sörbom, D. (1986). LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood. Morresville, IN: Scientific Software.

Jöreskog, K. G., & Sörbom, D. (1989). LISREL 7: A guide to the program and applications (2nd ed.). Chicago: SPSS.

Kalbfleisch, J. D., & Prentice, R. L. (1980). The statistical analysis of failure time data. New York: John Wiley and Sons.

Kieckhaefer, W. F., Ward, D. G., Kusulas, J. W., Cole, D. R., Rupp, L. M., & May, M. H. (1992). Criterion development for 18 technical training schools in the Navy, Army, and Air Force (Contract # N66001-90-D-9502, Delivery Order 7J08). San Diego: Navy Personnel Research and Development Center.

Klein, S., Hawes-Dawson, J., & Martin, T. (1991). Why recruits separate early (R-3980-FMP). Santa Monica, CA: Rand Corporation.

Knapp, D. J. (1993, August). Alternative conceptualizations of turnover. In J.P. Campbell (Chair), Prediction of turnover in a longitudinal sample using event history analysis. Symposium conducted at the convention of the American Psychological Association, Toronto.

LaRocco, J. M., Pugh, W. M., & Gunderson, E. K. (1977). Identifying determinants of retention decisions. Personnel Psychology, 30, 199-215.

Laurence, J. H. (1993). Education standards and military selection: From the beginning. In T. Trent & J. H. Laurence (Eds.), Adaptability screening for the Armed Forces (pp. 1-40). Washington, DC: Office of the Assistant Secretary of Defense.

Lawless, J. F. (1982). Statistical models and methods for lifetime data. New York: John Wiley and Sons.

Linn, R. L. (1968). Range restriction problems in the use of self selected groups for test validation. Psychological Bulletin, 69, 69-73.

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading MA: Addison-Wesley.

McBride, J. R. (1993). Compensatory screening model development. In T. Trent & J. H. Laurence (Eds.), Adaptability screening for the Armed Forces (pp. 1-40). Washington, DC: Office of the Assistant Secretary of Defense.

McCloy, R. A. (1993). An overview of survival analysis. Paper presented at the conference of the American Psychological Association, Toronto.

McCloy, R. A., Harris, D. A., Barnes, J. D., Hogan, P. F., Smith, D. A., Clifton, D., & Sola, M. (1992). Accession quality, job performance, and cost: A cost-performance tradeoff model (FR-PRD-92-11). Alexandria, VA: Human Resources Research Organization.

McCloy, R. A., & Oppler, S. H. (1990). End-of-training measures. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 fiscal year (ARI Technical Report 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 252 675)

McEvoy, G. M., & Cascio, W. F. (1987). Do good or poor performers leave? A meta-analysis of the relationship between performance and turnover. Academy of Management Journal, 30, 744-762.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. Personnal Psychology, 43, 335-354.

Miller, H. E., Katerberg, R., & Hulin, C. L. (1979). Evaluation of the Mobley, Horner, and Hollingsworth model of employee turnover. Journal of Applied Psychology, 64, 509-517.

Mobley, W. H., Hand, H. H., Baker, R. L., & Meglino, B. M. (1979). Conceptual and empirical analysis of military recruit training attrition. Journal of Applied Psychology, 64, 10-18.

Morita, J. G., Lee, T. W., & Mowday, R. T. (1989). Introducing survival analysis to organizational researchers: A selected application to turnover research. Journal of Applied Psychology, 74, 280-292.

Motowidlo, S. J., Dowell, B. E., Hopp, M. A., Borman, W. C., Johnson, P. D., & Dunnette, M. S. (1976). Motivation, satisfaction, and morale in Army careers: A review of theory and measurement (ARI Technical Report 76-A7). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Motowidlo, S. J., & Lawton, G. W. (1984). Affective and cognitive factors in soldiers' reenlistment decisions. Journal of Applied Psychology, 69, 157-166.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. Journal of Abnormal and Social Psychology, 66, 574-583.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw-Hill.

Oppler, S. H., Childs, R. A., & Peterson, N. G. (1994). Development of the longitudinal validation sample first-tour performance model. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1991 fiscal year (ARI Research Note 94-10). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 278 726)

Oppler, S. H, McCloy, R. A., & Peterson, N. G. (1994). Prediction of performance in training. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1991 fiscal year, Chapter 3 (ARI Research Note 94-10). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 278 726)

Oppler, S. H., Peterson, N. G., & Russell, T. (1994). Basic validation results for the LVI sample. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual report, 1991 fiscal year (ARI Research Note 94-10). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 278 726)

Organ, D. W. (1988). Organizational citizenship behavior: The good soldier syndrome. Lexington, MA: D. C. Heath.

Peterson, N. G., Hough, L. M., Dunnette, M. D., Rosse, R. L., Houston, J. S., Toquam, J. L., & Wing, H. (1990). Project A: Specification of the predictor domain and development of new selection/classification tests. Personnel Psychology, 43, 247-276.

Peterson, N. G., Oppler, S. H., Sager, C. E., & Rosse, R. L. (1993). Analysis of the Enhanced Computer Administered Test Battery: An evaluation of potential revisions and additions to the Armed Services Vocational Aptitude Battery (Draft report prepared for the Selection and Classification for Critical MOS Project). Washington, DC: American Institutes for Research. (ADA 252 675)

Peterson, N. G., Russell, T. L., Hallam, G., Hough, L. M., Owens-Kurtz, C. K., Gialluca, K., and Kerwin, C. (1990). Analysis of the Experimental Predictor Battery: LV Sample. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 fiscal year, Chapter 3 (ARI Technical Report 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 252 675)

Peterson, N., Russell, T., Hallam, G., Hough, L., Owens-Kurtz, C., Gialluca, K., & Kerwin, K. (1990). Analysis of the experimental predictor battery: LV sample. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual report, 1990 fiscal year (ARI Technical Report 952, pp. 73-199). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 252 675)

Petty, M. M., McGee, G. W., & Cavender, J. W. (1984). A meta-analysis of the relationships between individual job satisfaction and individual performance. Academy of Management Review, 9, 712-723.

Podsakoff, P. M., & Williams, L. (1986). The relationship between job performance and job satisfaction. In E. A. Locke (Ed.), Generalizing from laboratory to field settings (pp. 207-253). Lexington, MA: Lexington Books.

Ree, J. R., & Earles, J. A. (1991). Predicting training success: Not much more than g. Personnel Psychology, 44, 321-332.

Research Triangle Institute (1988). "Family research program and the Project A data base" (RTI/3795-44 WP). Draft working paper. Research Triangle Park, NC: Author.

Royle, M. H., & Robertson, D. W. (1980). Job satisfaction measures as predictors of retention for Navy enlisted personnel (TR 81-2). San Diego, CA: Navy Personnel Research and Development Center.

Rozeboom, W. W. (1978). Estimation of cross-validated multiple correlation: A clarification. Psychological Bulletin, 85, 1348-1351.

SAS Institute (1992). The PHREG procedure. In SAS Technical Report P-229, SAS/STAT software: Changes and enhancements, release 6.07 (pp. 433-479). Cary, NC: Author.

Scholarios, T. M. (1990). Maximizing potential classification efficiency: Selection of predictor measures based on alternative psychometric indices. Unpublished doctoral dissertation, George Washington University.

Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. Psychological Bulletin, 110(2), 268-290.

Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. Journal of Educational Statistics, 18(2), 155-195.

Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. Journal of Applied Psychology, 68(4), 653-663.

Smith, P.C., Kendall, L.M., & Hulin, C.L. (1969). Measurement of satisfaction in work and retirement. New York: Rand McNally.

Statman, M. A. (1993). Improving the effectiveness of employment testing through classification: Alternative methods of developing test composites for optimal job assignment and vocational counseling. Submitted doctoral dissertation, George Washington University.

Stein, C. (1960). Multiple regression. In I. Olkin (Ed.), Contributions to probability and statistics. Stanford, CA: Stanford University Press.

Steinhaus, S. D. (1988). Predicting military attrition from educational and biographical information (FR-PRD-88-06). Alexandria, VA: Human Resources Research Organization.

Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analyses based on meta-analytic findings. Personnel Psychology, 46, 259-293.

Trent, T. (1993). The Armed Services Applicant Profile (ASAP). In T. Trent and J. H. Laurence (Eds.), Adaptability screening for the Armed Forces (pp. 71-99). Washington, DC: Office of the Assistant Secretary of Defense.

Vroom, V. H. (1964). Work and motivation. New York: John Wiley & Sons.

Waters, B. K., Barnes, J. D., Foley, P., Steinhaus, S. D., & Brown, D. C. (1988). Estimating the reading skills of military applicants: Development of an ASVAB to RGL conversion table (FR-PRD-88-22). Alexandria, VA: Human Resources Research Organization.

Weiss, D. J., Dawis, R. V., England, G. W., & Lofquist, L. H. (1967). Manual for the Minnesota Satisfaction Questionnaire. Minnesota Studies in Vocational Rehabilitation, 22.

White, L. A. (1994). Development of composite scores for Assessment of the Background and Life Experiences (ABLE) Instrument. In J. P. Campbell & L. M. Zook (Eds.), Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1992 fiscal year (ARI Research Note 94-27, pp. 25-32). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (AD B188 259)

White, L. A., Nord, R. D., Mael, F. A., & Young, M. C. (1993). The Assessment of Background and Life Experiences (ABLE). In T. Trent & J. H. Laurence (Eds.), Adaptability screening for the Armed Forces (pp. 101-162). Washington, DC: Office of the Assistant Secretary of Defense.

Willett, J. B., & Singer, J. D. (1993). It's deja-vu all over again: Using multiple-spell discrete-time survival analysis. Manuscript submitted for publication.

Willett, J. B., & Singer, J. D. (in press). Investigating onset, cessation, relapse, and recovery: Why you should, and how you can, use discrete-time survival analysis. Journal of Consulting and Clinical Psychology.

Wise, L. L., McHenry, J. J., & Campbell, J. P. (1990). Identifying optimal predictor composites and testing for generalizability across jobs and performance factors. Personnel Psychology 43, 355-366.

Wise, L. L., McHenry, J. J., & Young, W. Y. (1986). Project A Concurrent Validation: Treatment of Missing Data. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wise, L. L., & McLaughlin, D. H. (1980). Guidebook for the imputation of missing data. Palo Alto, CA: American Institutes for Research.

Wise, L. L., Peterson, N. G., Hoffman, R. G., Campbell, J. P., & Arabian, J. M. (1991). The Army Synthetic Validity Project: Report of Phase III results, Vol. I (ARI Technical Report 922). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A235 635)

Wise, L. L., Welsh, J., Grafton, F., Foley, P., Earles, J., Swain, L., & Divgi, D. R., (1992). Sensitivity and fairness of the Armed Services Vocational Aptitude Battery (ASVAB) technical composites. Monterey CA: Defense Manpower Data Center, Personnel Testing Division.

Wolfe, J. H., Alderton, D. L., Larson, G. E., & Held, J. D. (in preparation). "Incremental validity of Enhanced Computer Administered Testing (ECAT)" (Draft Technical Report). San Diego, CA: Navy Personnel Research and Developmental Center.

Young, M. C., White, L. A., & Oppler, S. H. (1992). Effects of coaching on validity of a self-report temperament measure. Proceedings of the 34th Annual Conferences of the Military Testing Association (pp. 188-193). San Diego.

Youngblood, S. A., Mobley, W. H., & Meglino, B. M. (1983). A longitudinal analysis of the turnover process. Journal of Applied Psychology, 68, 507-516.

Appendix A

VALIDITIES OF THE TOP 20 POTENTIAL TEST BATTERIES RANKED ACCORDING TO EACH PERFORMANCE INDEX WITHIN EACH CRITERION AND EACH TIME INTERVAL

Note: For each battery the tests are listed in an arbitrary order (i.e., the order provides no information about the relative contribution of each test).

All variance/covariance matrices used to calculate validity coefficients are corrected for multivariate range restriction (Lord & Novick, 1968, p.147) before averaging.

Each validity estimate is adjusted using Rozeboom formula #8 before averaging.

Means are computed as unweighted (i.e., equally weighted) means across nine MOS. Corresponding SDs are "unbiased" estimates, based on the same nine MOS.

Discriminant validity is the difference between the unweighted mean validity and the mean correlation of the MOS-specific equations with the other eight MOS.

Brogden index of classification efficiency is the mean absolute validity times the square-root of (one minus the mean correlation among the predicted scores).

All test battery times are computed with 3 minutes added for each test, plus 20 minutes of start-up time.

Subgroup difference estimates = difference between mean composite of the majority group and the mean composite of the minority group divided by the majority group's standard deviation; these estimates are based on means and variances that have been corrected for range restriction (the means, variances, and covariances for the 10 ASVAB scores for majority and minority groups were obtained from the 1980 Youth Population and the remaining variances, covariances, and means are obtained by range restriction corrections).

Predictor Test Names

ASAR	Arithmetic Reasoning
ASWK	Word Knowledge
ASAS	Auto & Shop Information
ASCS	Coding Speed
ASEI	Electronics Information
ASGS	General Science
ASMC	Mechanical Comprehension
ASMK	Mathematical Knowledge
ASNO	Numerical Operations
ASPC	Paragraph Comprehension
SPAO	Assembling Objects
SPOR	Orientation
SPRS	Reasoning
CMST	Short-Term Memory
CMTI	Target Identification
CM1T	One-Hand Tracking
CM2T	Two-Hand Tracking

Appendix Table Headings

Validity Estimates
Discriminant Validity
Brogden Index
White-Black Subgroup Differences
White-Hispanic Subgroup Differences
Male-Female Subgroup Differences

Table A1.

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P D I F F E R E N C E--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.760	.052	.028	.149	1.529	0.181	1.143	0.118	0.354	0.730	102	4	ASAR,ASWK,ASAS,SPRS
2	.759	.053	.030	.156	1.489	0.183	1.082	0.111	0.424	0.834	102	5	ASAR,ASWK,ASAS,CMST,CM1T
3	.759	.051	.029	.153	1.448	0.227	1.037	0.156	0.356	0.756	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
4	.759	.053	.030	.154	1.494	0.202	1.086	0.122	0.407	0.849	102	5	ASAR,ASWK,ASAS,CMST,CM2T
5	.758	.052	.031	.157	1.493	0.245	1.090	0.187	0.374	0.765	101	5	ASAR,ASWK,ASAS,CMST,CMTI
6	.758	.053	.031	.156	1.535	0.141	1.096	0.097	0.417	0.861	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
7	.758	.052	.029	.152	1.452	0.202	1.063	0.125	0.379	0.728	101	5	ASAR,ASWK,ASAS,ASNO,CMST
8	.758	.053	.029	.151	1.483	0.172	1.106	0.112	0.389	0.719	101	4	ASAR,ASWK,ASAS,ASGS
9	.758	.053	.030	.154	1.542	0.168	1.101	0.106	0.397	0.873	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
10	.757	.053	.028	.149	1.490	0.180	1.086	0.115	0.367	0.727	95	4	ASAR,ASWK,ASAS,CMST
11	.757	.053	.028	.150	1.496	0.179	1.083	0.118	0.389	0.718	100	4	ASAR,ASWK,ASAS,SPOR
12	.757	.052	.031	.158	1.545	0.211	1.112	0.178	0.359	0.805	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
13	.757	.053	.029	.151	1.507	0.180	1.105	0.112	0.348	0.752	103	4	ASAR,ASWK,ASAS,ASPC
14	.756	.055	.032	.160	1.506	0.208	1.092	0.147	0.458	0.813	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
15	.756	.055	.032	.159	1.514	0.230	1.099	0.163	0.439	0.828	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
16	.756	.054	.031	.157	1.475	0.189	1.069	0.113	0.466	0.813	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
17	.756	.052	.028	.150	1.539	0.164	1.102	0.109	0.348	0.771	97	4	ASAR,ASWK,ASAS,ASCS
18	.756	.055	.031	.155	1.503	0.178	1.085	0.109	0.459	0.824	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
19	.755	.055	.030	.154	1.501	0.166	1.084	0.103	0.458	0.812	94	4	ASAR,ASWK,ASAS,CM1T
20	.755	.054	.031	.156	1.482	0.214	1.074	0.124	0.446	0.826	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A1. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Index	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.756	.055	.032	.160	1.506	0.208	1.092	0.147	0.458	0.813	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
2	.756	.055	.032	.159	1.514	0.230	1.099	0.163	0.439	0.828	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
3	.755	.054	.031	.158	1.487	0.233	1.088	0.181	0.407	0.755	99	5	ASAR,ASWK,ASAS,ASNO,CMTI
4	.757	.052	.031	.158	1.545	0.211	1.112	0.178	0.359	0.805	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
5	.756	.054	.031	.157	1.475	0.189	1.069	0.113	0.466	0.813	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
6	.758	.052	.031	.157	1.493	0.245	1.090	0.187	0.374	0.765	101	5	ASAR,ASWK,ASAS,CMST,CMTI
7	.755	.054	.031	.156	1.482	0.214	1.074	0.124	0.446	0.826	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
8	.758	.053	.031	.156	1.535	0.141	1.096	0.097	0.417	0.861	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
9	.754	.054	.031	.156	1.514	0.233	1.105	0.181	0.398	0.748	93	4	ASAR,ASWK,ASAS,CMTI
10	.756	.055	.031	.155	1.503	0.178	1.085	0.109	0.459	0.824	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
11	.759	.053	.030	.156	1.489	0.183	1.082	0.111	0.424	0.834	102	5	ASAR,ASWK,ASAS,CMST,CM1T
12	.755	.055	.030	.154	1.501	0.166	1.084	0.103	0.458	0.812	94	4	ASAR,ASWK,ASAS,CM1T
13	.758	.053	.030	.154	1.542	0.168	1.101	0.106	0.397	0.873	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
14	.759	.053	.030	.154	1.494	0.202	1.086	0.122	0.407	0.849	102	5	ASAR,ASWK,ASAS,CMST,CM2T
15	.755	.054	.030	.153	1.508	0.189	1.090	0.113	0.438	0.826	94	4	ASAR,ASWK,ASAS,CM2T
16	.759	.051	.029	.153	1.448	0.227	1.037	0.156	0.356	0.756	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
17	.758	.053	.029	.151	1.483	0.172	1.106	0.112	0.389	0.719	101	4	ASAR,ASWK,ASAS,ASGS
18	.758	.052	.029	.152	1.452	0.202	1.063	0.125	0.379	0.728	101	5	ASAR,ASWK,ASAS,ASNO,CMST
19	.757	.053	.029	.151	1.507	0.180	1.105	0.112	0.348	0.752	103	4	ASAR,ASWK,ASAS,ASPC
20	.756	.052	.028	.150	1.539	0.164	1.102	0.109	0.348	0.771	97	4	ASAR,ASWK,ASAS,ASCS
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A1. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.756	.055	.032	.160	1.506	0.208	1.092	0.147	0.458	0.813	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
2	.756	.055	.032	.159	1.514	0.230	1.099	0.163	0.439	0.828	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
3	.755	.054	.031	.158	1.487	0.233	1.088	0.181	0.407	0.755	99	5	ASAR,ASWK,ASAS,ASNO,CMTI
4	.757	.052	.031	.158	1.545	0.211	1.112	0.178	0.359	0.805	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
5	.758	.052	.031	.157	1.493	0.245	1.090	0.187	0.374	0.765	101	5	ASAR,ASWK,ASAS,CMST,CMTI
6	.756	.054	.031	.157	1.475	0.189	1.069	0.113	0.466	0.813	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
7	.758	.053	.031	.156	1.535	0.141	1.096	0.097	0.417	0.861	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
8	.755	.054	.031	.156	1.482	0.214	1.074	0.124	0.446	0.826	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
9	.754	.054	.031	.156	1.514	0.233	1.105	0.181	0.398	0.748	93	4	ASAR,ASWK,ASAS,CMTI
10	.759	.053	.030	.156	1.489	0.183	1.082	0.111	0.424	0.834	102	5	ASAR,ASWK,ASAS,CMST,CM1T
11	.756	.055	.031	.155	1.503	0.178	1.085	0.109	0.459	0.824	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
12	.759	.053	.030	.154	1.494	0.202	1.086	0.122	0.407	0.849	102	5	ASAR,ASWK,ASAS,CMST,CM2T
13	.758	.053	.030	.154	1.542	0.168	1.101	0.106	0.397	0.873	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
14	.755	.055	.030	.154	1.501	0.166	1.084	0.103	0.458	0.812	94	4	ASAR,ASWK,ASAS,CM1T
15	.759	.051	.029	.153	1.448	0.227	1.037	0.156	0.356	0.756	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
16	.755	.054	.030	.153	1.508	0.189	1.090	0.113	0.438	0.826	94	4	ASAR,ASWK,ASAS,CM2T
17	.758	.052	.029	.152	1.452	0.202	1.063	0.125	0.379	0.728	101	5	ASAR,ASWK,ASAS,ASNO,CMST
18	.758	.053	.029	.151	1.483	0.172	1.106	0.112	0.389	0.719	101	4	ASAR,ASWK,ASAS,ASGS
19	.757	.053	.029	.151	1.507	0.180	1.105	0.112	0.348	0.752	103	4	ASAR,ASWK,ASAS,ASPC
20	.756	.052	.028	.150	1.539	0.164	1.102	0.109	0.348	0.771	97	4	ASAR,ASWK,ASAS,ASCS
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A1. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Validty Mean	Est. S.D.	Discr Vldty	Brog. Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Test Time	NV	Predictor Variables
1	.736	.040	.006	.079	1.285	0.301	0.939	0.198	0.166	0.210	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.738	.039	.011	.100	1.295	0.237	0.951	0.132	0.181	0.298	103	6	ASAR,ASWK,ASCS,ASNO,CMST,CMTI
3	.733	.040	.006	.076	1.299	0.312	0.942	0.204	0.172	0.210	89	4	ASAR,ASWK,ASCS,ASNO
4	.740	.040	.010	.097	1.300	0.281	0.946	0.186	0.251	0.430	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
5	.734	.042	.007	.081	1.303	0.276	0.979	0.199	0.162	0.260	103	5	ASAR,ASWK,ASNO,ASPC,CMST
6	.730	.041	.006	.075	1.307	0.292	0.973	0.194	0.192	0.215	87	4	ASAR,ASWK,ASNO,CMST
7	.741	.039	.012	.103	1.309	0.241	0.951	0.164	0.244	0.504	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
8	.737	.041	.010	.094	1.316	0.290	0.951	0.191	0.272	0.408	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
9	.735	.040	.010	.097	1.317	0.254	0.963	0.141	0.194	0.285	95	5	ASAR,ASWK,ASCS,ASNO,CMTI
10	.733	.040	.011	.097	1.322	0.230	0.989	0.131	0.210	0.283	93	5	ASAR,ASWK,ASNO,CMST,CMTI
11	.730	.042	.006	.079	1.323	0.288	0.989	0.206	0.170	0.261	95	4	ASAR,ASWK,ASNO,ASPC
12	.735	.041	.010	.095	1.323	0.271	0.978	0.184	0.289	0.419	94	5	ASAR,ASWK,ASNO,CMST,CM1T
13	.739	.041	.013	.106	1.324	0.240	0.961	0.131	0.276	0.434	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM1T
14	.739	.040	.012	.102	1.324	0.248	0.955	0.166	0.280	0.475	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
15	.738	.039	.011	.099	1.327	0.250	0.957	0.168	0.262	0.482	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
16	.736	.041	.013	.108	1.329	0.216	0.985	0.121	0.292	0.451	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM1T
17	.725	.042	.006	.071	1.331	0.306	0.984	0.203	0.205	0.216	79	3	ASAR,ASWK,ASNO
18	.736	.040	.012	.104	1.331	0.232	0.982	0.162	0.297	0.485	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
19	.736	.040	.012	.101	1.332	0.234	0.983	0.165	0.281	0.489	94	5	ASAR,ASWK,ASNO,CMST,CM2T
20	.739	.039	.014	.110	1.334	0.213	0.967	0.138	0.267	0.499	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A1. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.736	.040	.006	.079	1.285	0.301	0.939	0.198	0.166	0.210	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.733	.040	.006	.076	1.299	0.312	0.942	0.204	0.172	0.210	89	4	ASAR,ASWK,ASCS,ASNO
3	.740	.040	.010	.097	1.300	0.281	0.946	0.186	0.251	0.430	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
4	.737	.041	.010	.094	1.316	0.290	0.951	0.191	0.272	0.408	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
5	.741	.039	.012	.103	1.309	0.241	0.951	0.164	0.244	0.504	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
6	.738	.039	.011	.100	1.295	0.237	0.951	0.132	0.181	0.298	103	6	ASAR,ASWK,ASCS,ASNO,CMST,CMTI
7	.739	.040	.012	.102	1.324	0.248	0.955	0.166	0.280	0.475	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
8	.738	.039	.011	.099	1.327	0.250	0.957	0.168	0.262	0.482	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
9	.739	.041	.013	.106	1.324	0.240	0.961	0.131	0.276	0.434	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM1T
10	.735	.040	.010	.097	1.317	0.254	0.963	0.141	0.194	0.285	95	5	ASAR,ASWK,ASCS,ASNO,CMTI
11	.740	.040	.009	.091	1.335	0.295	0.965	0.195	0.202	0.247	102	5	ASAR,ASWK,ASCS,ASNO,SPOR
12	.739	.039	.014	.110	1.334	0.213	0.967	0.138	0.267	0.499	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM2T
13	.730	.041	.006	.075	1.307	0.292	0.973	0.194	0.192	0.215	87	4	ASAR,ASWK,ASNO,CMST
14	.735	.041	.010	.095	1.323	0.271	0.978	0.184	0.289	0.419	94	5	ASAR,ASWK,ASNO,CMST,CM1T
15	.734	.042	.007	.081	1.303	0.276	0.979	0.199	0.162	0.260	103	5	ASAR,ASWK,ASNO,ASPC,CMST
16	.736	.040	.012	.104	1.331	0.232	0.982	0.162	0.297	0.485	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
17	.736	.040	.012	.101	1.332	0.234	0.983	0.165	0.281	0.489	94	5	ASAR,ASWK,ASNO,CMST,CM2T
18	.725	.042	.006	.071	1.331	0.306	0.984	0.203	0.205	0.216	79	3	ASAR,ASWK,ASNO
19	.736	.041	.013	.108	1.329	0.216	0.985	0.121	0.292	0.451	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM1T
20	.731	.043	.009	.090	1.347	0.285	0.989	0.191	0.321	0.389	86	4	ASAR,ASWK,ASNO,CM1T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A1. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.724	.043	.004	.066	1.426	0.095	1.050	0.060	0.081	0.183	99	4	ASAR,ASWK,ASCS,ASPC
2	.726	.044	.002	.056	1.405	0.076	1.044	0.051	0.084	0.105	97	4	ASAR,ASWK,ASPC,CMST
3	.733	.043	.003	.061	1.452	0.065	1.112	0.043	0.085	0.094	104	4	ASAR,ASWK,ASPC,SPRS
4	.731	.041	.004	.068	1.462	0.099	1.105	0.047	0.096	0.154	98	4	ASAR,ASWK,ASCS,SPRS
5	.723	.044	.002	.051	1.418	0.065	1.050	0.046	0.098	0.100	89	3	ASAR,ASWK,ASPC
6	.724	.042	.003	.062	1.412	0.103	1.036	0.052	0.098	0.157	91	4	ASAR,ASWK,ASCS,CMST
7	.721	.042	.003	.059	1.432	0.105	1.042	0.051	0.105	0.157	83	3	ASAR,ASWK,ASCS
8	.732	.042	.002	.053	1.441	0.076	1.096	0.050	0.106	0.039	96	4	ASAR,ASWK,SPRS,CMST
9	.729	.042	.008	.087	1.420	0.123	1.061	0.114	0.108	0.217	103	5	ASAR,ASWK,ASPC,CMST,CMTI
10	.738	.043	.005	.073	1.479	0.087	1.036	0.060	0.110	0.151	104	4	ASAR,ASWK,ASCS,SPAO
11	.733	.040	.009	.093	1.467	0.067	1.113	0.099	0.113	0.278	104	5	ASAR,ASWK,ASCS,SPRS,CMTI
12	.738	.044	.003	.060	1.461	0.063	1.033	0.063	0.115	0.042	102	4	ASAR,ASWK,SPAO,CMST
13	.723	.043	.001	.042	1.405	0.076	1.035	0.051	0.115	0.045	81	3	ASAR,ASWK,CMST
14	.734	.041	.007	.084	1.444	0.125	1.101	0.118	0.117	0.165	102	5	ASAR,ASWK,SPRS,CMST,CMTI
15	.730	.043	.002	.048	1.454	0.062	1.106	0.040	0.118	0.043	88	3	ASAR,ASWK,SPRS
16	.726	.042	.004	.065	1.362	0.118	1.020	0.056	0.122	0.052	100	3	ASAR,ASWK,ASMK
17	.727	.041	.009	.090	1.426	0.087	1.053	0.110	0.123	0.286	97	5	ASAR,ASWK,ASCS,CMST,CMTI
18	.737	.044	.003	.054	1.475	0.050	1.036	0.060	0.126	0.036	94	3	ASAR,ASWK,SPAO
19	.726	.043	.007	.082	1.439	0.094	1.075	0.094	0.128	0.186	95	4	ASAR,ASWK,ASPC,CMTI
20	.732	.042	.006	.079	1.461	0.092	1.115	0.094	0.134	0.131	94	4	ASAR,ASWK,SPRS,CMTI
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

Table A2.

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.775	.048	.034	.168	1.449	0.314	1.064	0.214	0.332	0.772	163	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CMTI
2	.774	.049	.033	.166	1.458	0.248	1.074	0.141	0.369	0.838	164	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CM1T
3	.774	.049	.033	.165	1.460	0.272	1.075	0.148	0.356	0.857	164	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CM2T
4	.774	.051	.033	.166	1.491	0.238	1.090	0.183	0.310	0.833	164	9	ASAR,ASWK,ASAS,ASEI,ASGS,ASPC,SPAO,CMST,CMTI
5	.774	.048	.035	.171	1.477	0.299	1.067	0.191	0.354	0.871	164	8	ASAR,ASWK,ASAS,ASMK,ASPC,SPAO,CMTI,CM1T
6	.774	.051	.033	.168	1.492	0.233	1.083	0.162	0.333	0.892	159	9	ASAR,ASWK,ASAS,ASGS,ASPC,SPAO,CMST,CMTI,CM1T
7	.774	.048	.035	.170	1.481	0.315	1.070	0.203	0.340	0.884	164	8	ASAR,ASWK,ASAS,ASMK,ASPC,SPAO,CMTI,CM2T
8	.774	.049	.033	.167	1.453	0.311	1.057	0.219	0.346	0.799	161	8	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMST,CMTI
9	.774	.051	.033	.166	1.495	0.252	1.086	0.174	0.320	0.908	159	9	ASAR,ASWK,ASAS,ASGS,ASPC,SPAO,CMST,CMTI,CM2T
10	.773	.050	.033	.166	1.510	0.219	1.092	0.174	0.286	0.838	162	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASPC,SPAO,CMST,CMTI
11	.773	.050	.033	.165	1.461	0.244	1.066	0.141	0.382	0.868	162	8	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMST,CM1T
12	.773	.049	.031	.161	1.475	0.257	1.075	0.136	0.308	0.737	159	7	ASAR,ASWK,ASAS,ASMK,ASPC,SPAO,CMST
13	.773	.050	.032	.163	1.480	0.243	1.083	0.140	0.328	0.771	163	7	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,SPAO
14	.773	.050	.033	.164	1.462	0.266	1.068	0.148	0.369	0.882	162	8	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMST,CM2T
15	.773	.051	.031	.162	1.476	0.233	1.072	0.166	0.302	0.768	162	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,ASPC,SPAO,CMST
16	.773	.050	.033	.168	1.470	0.304	1.058	0.244	0.322	0.831	164	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASNO,SPAO,CMST,CMTI
17	.773	.049	.031	.162	1.469	0.244	1.082	0.141	0.334	0.740	163	8	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,CMST
18	.773	.049	.034	.169	1.462	0.286	1.048	0.213	0.345	0.890	159	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,SPAO,CMST,CMTI,CM1T
19	.773	.049	.035	.170	1.461	0.288	1.065	0.190	0.385	0.843	162	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMTI,CM1T
20	.773	.049	.031	.161	1.460	0.248	1.076	0.141	0.339	0.717	157	7	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A2. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

	Validty Est.		Discr	Brog.	--SUBGROUP DIFFERENCE-- Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.772	.049	.036	.174	1.433	0.297	1.078	0.201	0.403	0.892	163	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMST,CMTI,CM1T
2	.770	.049	.036	.174	1.450	0.268	1.098	0.173	0.411	0.861	163	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,CMTI,CM1T
3	.769	.050	.036	.174	1.458	0.273	1.095	0.184	0.423	0.882	161	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASNO,ASPC,CMTI,CM1T
4	.768	.051	.036	.173	1.452	0.272	1.097	0.178	0.445	0.859	159	9	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,CMTI,CM1T
5	.770	.048	.036	.173	1.433	0.292	1.071	0.187	0.393	0.879	157	9	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
6	.770	.048	.036	.173	1.434	0.300	1.071	0.190	0.394	0.886	164	10	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
7	.770	.050	.036	.173	1.463	0.252	1.092	0.181	0.424	0.882	163	9	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,CMTI,CM1T
8	.769	.050	.036	.172	1.448	0.280	1.088	0.191	0.429	0.872	155	8	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMTI,CM1T
9	.771	.049	.036	.173	1.418	0.289	1.074	0.194	0.426	0.871	161	9	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,CMST,CMTI,CM1T
10	.770	.050	.036	.172	1.444	0.287	1.093	0.186	0.416	0.852	164	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASPC,CMTI,CM1T,CM2T
11	.770	.049	.036	.172	1.443	0.281	1.093	0.185	0.416	0.843	157	8	ASAR,ASWK,ASAS,ASGS,ASMK,ASPC,CMTI,CM1T
12	.769	.050	.036	.172	1.449	0.285	1.089	0.191	0.429	0.879	162	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMTI,CM1T,CM2T
13	.771	.047	.036	.173	1.453	0.286	1.077	0.192	0.378	0.890	161	9	ASAR,ASWK,ASAS,ASCS,ASMK,ASPC,CMST,CMTI,CM1T
14	.770	.049	.036	.173	1.429	0.301	1.077	0.187	0.412	0.868	162	10	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
15	.770	.049	.036	.173	1.428	0.292	1.077	0.184	0.412	0.859	155	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T
16	.772	.049	.036	.172	1.437	0.315	1.081	0.215	0.390	0.897	163	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMST,CMTI,CM2T
17	.770	.048	.036	.172	1.439	0.313	1.074	0.198	0.392	0.877	158	9	ASAR,ASWK,ASAS,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
18	.770	.048	.036	.172	1.438	0.307	1.074	0.196	0.391	0.868	151	8	ASAR,ASWK,ASAS,ASMK,ASPC,CMST,CMTI,CM1T
19	.771	.048	.036	.173	1.444	0.275	1.079	0.187	0.394	0.873	159	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,CMST,CMTI,CM1T
20	.769	.050	.036	.172	1.464	0.290	1.100	0.199	0.407	0.886	161	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASNO,ASPC,CMTI,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A2. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

	Validty Est.		Discr	Brog.	--S U B G R O U P	D I F F E R E N C E--					Test		
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.772	.049	.036	.174	1.433	0.297	1.078	0.201	0.403	0.892	163	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMST,CMTI,CM1T
2	.769	.050	.036	.174	1.458	0.273	1.095	0.184	0.423	0.882	161	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASNO,ASPC,CMTI,CM1T
3	.770	.049	.036	.174	1.450	0.268	1.098	0.173	0.411	0.861	163	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,CMTI,CM1T
4	.768	.051	.036	.173	1.452	0.272	1.097	0.178	0.445	0.859	159	9	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,CMTI,CM1T
5	.770	.050	.036	.173	1.463	0.252	1.092	0.181	0.424	0.882	163	9	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,CMTI,CM1T
6	.770	.048	.036	.173	1.434	0.300	1.071	0.190	0.394	0.886	164	10	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
7	.770	.048	.036	.173	1.433	0.292	1.071	0.187	0.393	0.879	157	9	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
8	.771	.049	.036	.173	1.418	0.289	1.074	0.194	0.426	0.871	161	9	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,CMST,CMTI,CM1T
9	.770	.049	.036	.173	1.439	0.298	1.063	0.211	0.409	0.894	163	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMK,ASNO,CMST,CMTI,CM1T
10	.771	.047	.036	.173	1.453	0.286	1.077	0.192	0.378	0.890	161	9	ASAR,ASWK,ASAS,ASCS,ASMK,ASPC,CMST,CMTI,CM1T
11	.770	.049	.036	.173	1.429	0.301	1.077	0.187	0.412	0.868	162	10	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
12	.770	.049	.036	.173	1.428	0.292	1.077	0.184	0.412	0.859	155	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T
13	.769	.049	.036	.173	1.436	0.305	1.073	0.198	0.426	0.885	160	10	ASAR,ASWK,ASAS,ASEI,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
14	.771	.048	.036	.173	1.444	0.275	1.079	0.187	0.394	0.873	159	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,CMST,CMTI,CM1T
15	.772	.048	.036	.173	1.438	0.313	1.070	0.199	0.384	0.869	164	9	ASAR,ASWK,ASAS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T
16	.769	.049	.036	.173	1.435	0.297	1.073	0.195	0.426	0.878	153	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASNO,CMST,CMTI,CM1T
17	.769	.049	.036	.173	1.457	0.273	1.082	0.175	0.403	0.868	162	9	ASAR,ASWK,ASAS,ASMK,ASNO,ASPC,SPOR,CMTI,CM1T
18	.769	.050	.036	.172	1.449	0.285	1.089	0.191	0.429	0.879	162	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMTI,CM1T,CM2T
19	.772	.049	.036	.172	1.437	0.315	1.081	0.215	0.390	0.897	163	9	ASAR,ASWK,ASAS,ASEI,ASMK,ASPC,CMST,CMTI,CM2T
20	.770	.050	.036	.172	1.444	0.287	1.093	0.186	0.416	0.852	164	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASPC,CMTI,CM1T,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A2. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.741	.039	.012	.107	1.273	0.316	0.948	0.258	0.136	0.264	140	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.743	.038	.017	.124	1.283	0.255	0.959	0.196	0.151	0.370	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI
3	.744	.039	.016	.121	1.290	0.294	0.956	0.248	0.223	0.489	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
4	.746	.038	.019	.132	1.292	0.239	0.959	0.187	0.226	0.529	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
5	.744	.038	.019	.129	1.293	0.248	0.952	0.176	0.246	0.511	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T
6	.740	.038	.017	.123	1.295	0.213	0.989	0.156	0.162	0.391	136	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CMTI
7	.742	.039	.016	.119	1.299	0.252	0.980	0.208	0.242	0.514	137	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T
8	.746	.038	.018	.126	1.300	0.250	0.962	0.221	0.233	0.548	154	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
9	.744	.037	.017	.122	1.300	0.259	0.954	0.210	0.252	0.533	138	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
10	.746	.037	.017	.124	1.301	0.253	0.963	0.223	0.219	0.545	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
11	.747	.038	.020	.135	1.301	0.229	0.964	0.169	0.235	0.573	160	10	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
12	.745	.037	.020	.132	1.302	0.241	0.957	0.157	0.254	0.556	144	9	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
13	.743	.039	.020	.131	1.303	0.211	0.985	0.147	0.245	0.554	143	8	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
14	.741	.038	.017	.122	1.303	0.269	0.972	0.205	0.161	0.363	138	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMTI
15	.747	.037	.020	.133	1.303	0.234	0.966	0.173	0.222	0.574	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM2T
16	.745	.037	.020	.130	1.304	0.245	0.959	0.162	0.241	0.558	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM2T
17	.743	.039	.016	.118	1.305	0.301	0.962	0.252	0.240	0.475	139	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
18	.747	.038	.015	.117	1.305	0.293	0.967	0.246	0.164	0.309	153	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST
19	.744	.038	.014	.113	1.306	0.301	0.960	0.235	0.184	0.291	137	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
20	.748	.037	.019	.131	1.308	0.241	0.971	0.191	0.171	0.399	159	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A2. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP	DIFFERENCE--							
	Validty Est.		Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.741	.039	.012	.107	1.273	0.316	0.948	0.258	0.136	0.264	140	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.744	.038	.019	.129	1.293	0.248	0.952	0.176	0.246	0.511	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T
3	.744	.037	.017	.122	1.300	0.259	0.954	0.210	0.252	0.533	138	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
4	.744	.039	.016	.121	1.290	0.294	0.956	0.248	0.223	0.489	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
5	.745	.037	.020	.132	1.302	0.241	0.957	0.157	0.254	0.556	144	9	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
6	.752	.038	.019	.131	1.347	0.231	0.958	0.165	0.205	0.514	158	9	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CMTI,CM1T
7	.750	.038	.017	.124	1.344	0.241	0.959	0.172	0.148	0.352	151	8	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CMTI
8	.745	.037	.020	.130	1.304	0.245	0.959	0.162	0.241	0.558	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM2T
9	.746	.038	.019	.132	1.292	0.239	0.959	0.187	0.226	0.529	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
10	.743	.038	.017	.124	1.283	0.255	0.959	0.196	0.151	0.370	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI
11	.750	.040	.015	.118	1.363	0.226	0.959	0.165	0.217	0.529	138	9	ASAR,ASWK,ASCS,ASNO,SPAO,CMST,CMTI,CM1T,CM2T
12	.744	.038	.014	.113	1.306	0.301	0.960	0.235	0.184	0.291	137	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
13	.743	.039	.016	.118	1.305	0.301	0.962	0.252	0.240	0.475	139	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
14	.746	.038	.018	.126	1.300	0.250	0.962	0.221	0.233	0.548	154	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
15	.752	.037	.020	.133	1.354	0.252	0.962	0.159	0.199	0.566	158	9	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CMTI,CM2T
16	.748	.037	.020	.135	1.313	0.236	0.962	0.171	0.250	0.521	150	9	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CMTI,CM1T
17	.748	.039	.012	.108	1.348	0.289	0.963	0.230	0.147	0.243	145	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST
18	.747	.038	.017	.125	1.315	0.286	0.963	0.228	0.248	0.486	144	8	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CM1T
19	.746	.037	.017	.124	1.301	0.253	0.963	0.223	0.219	0.545	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
20	.751	.040	.015	.120	1.364	0.222	0.963	0.140	0.220	0.489	144	9	ASAR,ASWK,ASCS,ASNO,SPAO,SPOR,CMST,CMTI,CM1T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A2. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.738	.041	.008	.093	1.394	0.123	1.079	0.076	0.072	0.173	149	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS,CMST
2	.738	.042	.007	.086	1.396	0.137	1.080	0.080	0.073	0.103	139	6	ASAR,ASWK,ASMK,ASPC,SPRS,CMST
3	.737	.041	.008	.091	1.409	0.110	1.088	0.066	0.075	0.175	141	6	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS
4	.731	.042	.007	.087	1.355	0.117	1.023	0.063	0.077	0.180	134	6	ASAR,ASWK,ASCS,ASMK,ASPC,CMST
5	.746	.042	.008	.093	1.431	0.158	1.055	0.077	0.077	0.101	160	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMST
6	.745	.042	.011	.104	1.463	0.093	1.066	0.092	0.079	0.307	149	8	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMST,CMTI
7	.745	.043	.009	.096	1.465	0.127	1.068	0.094	0.080	0.210	139	7	ASAR,ASWK,ASPC,SPAO,SPRS,CMST,CMTI
8	.743	.044	.006	.084	1.469	0.074	1.071	0.062	0.080	0.175	143	7	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMST
9	.745	.042	.008	.091	1.420	0.155	1.026	0.067	0.081	0.104	145	6	ASAR,ASWK,ASMK,ASPC,SPAO,CMST
10	.743	.044	.006	.082	1.481	0.079	1.077	0.063	0.082	0.175	135	6	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS
11	.746	.042	.010	.098	1.439	0.115	1.059	0.069	0.082	0.173	162	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPRS
12	.747	.041	.013	.112	1.417	0.263	1.024	0.154	0.082	0.230	151	7	ASAR,ASWK,ASMK,ASPC,SPAO,CMST,CMTI
13	.746	.042	.008	.090	1.440	0.137	1.060	0.069	0.083	0.095	152	6	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS
14	.745	.042	.010	.098	1.415	0.127	1.024	0.069	0.084	0.176	155	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,CMST
15	.740	.040	.013	.109	1.400	0.249	1.085	0.176	0.084	0.231	145	7	ASAR,ASWK,ASMK,ASPC,SPRS,CMST,CMTI
16	.740	.040	.014	.114	1.398	0.191	1.084	0.171	0.084	0.311	155	8	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS,CMST,CMTI
17	.745	.042	.011	.102	1.478	0.065	1.075	0.076	0.085	0.299	141	7	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMTI
18	.747	.041	.013	.109	1.438	0.227	1.059	0.132	0.085	0.203	158	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMTI
19	.747	.040	.015	.117	1.412	0.198	1.022	0.139	0.086	0.308	161	8	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,CMST,CMTI
20	.744	.042	.010	.096	1.428	0.114	1.028	0.066	0.086	0.176	147	6	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

Table A3.

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.779	.047	.036	.177	1.472	0.284	1.095	0.188	0.326	0.934	223	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM1T
2	.779	.047	.036	.176	1.475	0.301	1.098	0.201	0.315	0.943	223	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM2T
3	.779	.048	.036	.176	1.459	0.302	1.091	0.193	0.337	0.916	213	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM1T
4	.779	.048	.036	.177	1.468	0.299	1.098	0.190	0.332	0.923	219	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T
5	.779	.048	.036	.176	1.459	0.307	1.091	0.193	0.337	0.930	220	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
6	.779	.048	.036	.175	1.461	0.320	1.094	0.206	0.327	0.926	213	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM2T
7	.779	.048	.036	.176	1.471	0.317	1.101	0.204	0.322	0.930	219	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM2T
8	.779	.048	.036	.178	1.461	0.288	1.073	0.201	0.337	0.933	221	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
9	.779	.047	.035	.174	1.475	0.293	1.100	0.208	0.299	0.845	216	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI
10	.779	.048	.036	.176	1.463	0.305	1.075	0.213	0.325	0.943	221	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,CMST,CMTI,CM2T
11	.779	.048	.036	.177	1.464	0.286	1.074	0.200	0.337	0.929	208	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,CMST,CMTI,CM1T
12	.779	.048	.036	.177	1.448	0.306	1.069	0.205	0.348	0.916	211	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
13	.779	.048	.036	.177	1.463	0.291	1.074	0.200	0.336	0.941	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
14	.779	.048	.036	.178	1.458	0.304	1.075	0.204	0.343	0.922	217	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
15	.779	.049	.036	.177	1.447	0.312	1.068	0.207	0.348	0.929	218	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
16	.779	.047	.035	.175	1.473	0.305	1.098	0.220	0.300	0.845	222	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI
17	.779	.048	.035	.174	1.459	0.310	1.093	0.210	0.314	0.830	219	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
18	.779	.047	.036	.176	1.471	0.286	1.090	0.182	0.319	0.912	224	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
19	.779	.048	.036	.176	1.466	0.302	1.077	0.211	0.324	0.938	208	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,CMST,CMTI,CM2T
20	.779	.048	.036	.178	1.457	0.311	1.075	0.206	0.342	0.934	224	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A3. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P D I F F E R E N C E--								
	Validty Est.		Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.774	.049	.037	.177	1.438	0.292	1.090	0.197	0.391	0.886	205	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
2	.774	.049	.037	.177	1.444	0.275	1.087	0.194	0.381	0.901	209	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T
3	.774	.049	.037	.177	1.438	0.300	1.090	0.200	0.391	0.895	212	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
4	.774	.049	.037	.176	1.444	0.282	1.089	0.196	0.384	0.916	194	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
5	.774	.049	.037	.176	1.426	0.291	1.082	0.197	0.398	0.883	199	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T
6	.775	.049	.037	.177	1.435	0.299	1.087	0.201	0.387	0.899	196	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T
7	.774	.049	.037	.176	1.426	0.299	1.082	0.199	0.398	0.892	206	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
8	.774	.049	.037	.177	1.436	0.304	1.081	0.217	0.384	0.915	200	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
9	.774	.049	.037	.177	1.445	0.282	1.087	0.195	0.381	0.910	216	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
10	.774	.049	.037	.178	1.441	0.294	1.083	0.213	0.381	0.900	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
11	.775	.049	.037	.177	1.436	0.307	1.087	0.204	0.386	0.906	203	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
12	.775	.048	.036	.177	1.443	0.282	1.085	0.199	0.376	0.912	200	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T
13	.772	.050	.036	.176	1.464	0.254	1.098	0.184	0.400	0.892	201	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMTI,CM1T
14	.771	.050	.036	.176	1.458	0.269	1.104	0.182	0.412	0.873	197	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMTI,CM1T
15	.774	.049	.036	.178	1.441	0.302	1.083	0.216	0.381	0.909	222	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
16	.775	.049	.036	.176	1.426	0.307	1.080	0.205	0.391	0.902	197	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
17	.779	.048	.036	.177	1.464	0.286	1.074	0.200	0.337	0.929	208	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,CMST,CMTI,CM1T
18	.775	.049	.036	.177	1.443	0.289	1.085	0.201	0.376	0.920	207	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
19	.778	.048	.036	.177	1.449	0.302	1.069	0.202	0.348	0.911	220	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASPC,SPAO,CMST,CMTI,CM1T
20	.775	.049	.036	.178	1.439	0.298	1.087	0.201	0.384	0.891	218	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A3. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.777	.048	.036	.178	1.456	0.296	1.108	0.206	0.353	0.923	221	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
2	.779	.048	.036	.178	1.458	0.304	1.075	0.204	0.343	0.922	217	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
3	.779	.048	.036	.178	1.461	0.288	1.073	0.201	0.337	0.933	221	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
4	.775	.049	.036	.178	1.439	0.298	1.087	0.201	0.384	0.891	218	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T
5	.775	.048	.036	.178	1.438	0.299	1.079	0.216	0.377	0.910	206	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T
6	.774	.049	.037	.178	1.441	0.294	1.083	0.213	0.381	0.900	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
7	.779	.048	.036	.178	1.457	0.311	1.075	0.206	0.342	0.934	224	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
8	.777	.048	.036	.178	1.454	0.294	1.115	0.191	0.362	0.912	211	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
9	.779	.048	.036	.178	1.460	0.303	1.071	0.216	0.337	0.928	214	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T
10	.775	.049	.036	.178	1.444	0.281	1.083	0.198	0.375	0.905	222	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T
11	.777	.048	.036	.178	1.454	0.303	1.115	0.194	0.361	0.919	218	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
12	.774	.049	.036	.178	1.441	0.302	1.083	0.216	0.381	0.909	222	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
13	.775	.049	.036	.178	1.438	0.309	1.079	0.220	0.376	0.919	213	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
14	.776	.048	.036	.178	1.457	0.292	1.120	0.190	0.363	0.904	220	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,SPRS,CMST,CMTI,CM1T
15	.775	.049	.037	.177	1.435	0.299	1.087	0.201	0.387	0.899	196	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T
16	.778	.048	.036	.177	1.460	0.311	1.071	0.219	0.337	0.940	221	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
17	.777	.048	.036	.177	1.460	0.280	1.112	0.190	0.353	0.925	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
18	.777	.049	.036	.177	1.477	0.284	1.079	0.209	0.349	0.926	219	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMTI,CM1T
19	.775	.049	.037	.177	1.436	0.307	1.087	0.204	0.386	0.906	203	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
20	.774	.049	.037	.177	1.438	0.292	1.090	0.197	0.391	0.886	205	12	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A3. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.756	.038	.022	.143	1.355	0.237	0.972	0.162	0.198	0.603	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
2	.755	.038	.019	.135	1.356	0.242	0.991	0.198	0.136	0.414	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
3	.756	.038	.021	.142	1.358	0.234	0.988	0.193	0.182	0.563	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
4	.755	.038	.020	.138	1.361	0.233	0.991	0.170	0.179	0.584	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
5	.759	.039	.022	.144	1.361	0.216	1.047	0.159	0.233	0.591	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
6	.757	.038	.022	.144	1.362	0.235	0.991	0.178	0.187	0.609	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
7	.757	.038	.022	.142	1.364	0.241	0.992	0.185	0.179	0.613	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
8	.760	.039	.023	.146	1.366	0.238	1.049	0.158	0.238	0.631	202	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
9	.755	.039	.018	.132	1.367	0.278	0.998	0.245	0.186	0.507	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
10	.760	.038	.022	.144	1.367	0.244	1.050	0.164	0.230	0.640	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM2T
11	.756	.038	.022	.143	1.371	0.238	1.007	0.152	0.197	0.621	199	12	ASAR,ASWK,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
12	.756	.039	.020	.136	1.371	0.246	1.001	0.224	0.191	0.570	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
13	.756	.039	.019	.134	1.371	0.250	1.002	0.229	0.183	0.571	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
14	.756	.039	.021	.139	1.373	0.245	0.998	0.199	0.191	0.548	194	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T
15	.759	.040	.021	.139	1.374	0.220	1.058	0.188	0.241	0.603	196	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CM1T,CM2T
16	.762	.040	.022	.144	1.377	0.221	1.015	0.158	0.221	0.592	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
17	.761	.040	.020	.138	1.377	0.206	1.018	0.158	0.181	0.462	194	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI
18	.761	.039	.022	.142	1.378	0.242	1.014	0.168	0.219	0.619	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
19	.756	.038	.022	.142	1.378	0.219	1.001	0.183	0.196	0.594	201	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
20	.763	.039	.023	.146	1.380	0.244	1.016	0.169	0.224	0.637	208	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A3. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.756	.038	.022	.143	1.355	0.237	0.972	0.162	0.198	0.603	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
2	.756	.038	.021	.142	1.358	0.234	0.988	0.193	0.182	0.563	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
3	.757	.038	.022	.144	1.362	0.235	0.991	0.178	0.187	0.609	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
4	.755	.038	.020	.138	1.361	0.233	0.991	0.170	0.179	0.584	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
5	.755	.038	.019	.135	1.356	0.242	0.991	0.198	0.136	0.414	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
6	.757	.038	.022	.142	1.364	0.241	0.992	0.185	0.179	0.613	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
7	.756	.039	.021	.139	1.373	0.245	0.998	0.199	0.191	0.548	194	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T
8	.755	.039	.018	.132	1.367	0.278	0.998	0.245	0.186	0.507	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
9	.756	.038	.022	.142	1.378	0.219	1.001	0.183	0.196	0.594	201	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
10	.756	.039	.020	.136	1.371	0.246	1.001	0.224	0.191	0.570	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
11	.756	.039	.019	.134	1.371	0.250	1.002	0.229	0.183	0.571	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
12	.756	.038	.021	.140	1.380	0.226	1.003	0.191	0.186	0.600	194	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM2T
13	.756	.038	.022	.143	1.371	0.238	1.007	0.152	0.197	0.621	199	12	ASAR,ASWK,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
14	.760	.039	.025	.151	1.408	0.254	1.008	0.159	0.259	0.688	200	12	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
15	.755	.039	.019	.134	1.384	0.250	1.008	0.226	0.198	0.560	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CM1T,CM2T
16	.761	.039	.022	.142	1.378	0.242	1.014	0.168	0.219	0.619	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
17	.762	.039	.026	.153	1.406	0.225	1.014	0.151	0.237	0.676	209	12	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
18	.762	.039	.026	.154	1.407	0.242	1.014	0.160	0.239	0.706	216	13	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
19	.762	.040	.022	.144	1.377	0.221	1.015	0.158	0.221	0.592	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
20	.762	.039	.026	.152	1.408	0.223	1.015	0.150	0.236	0.668	196	11	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

(Continued)

Table A3. (Cont'd)

Technical/School Knowledge: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.755	.038	.019	.135	1.356	0.242	0.991	0.198	0.136	0.414	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
2	.753	.038	.020	.136	1.419	0.229	1.038	0.151	0.157	0.598	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
3	.760	.040	.018	.132	1.422	0.191	1.072	0.150	0.157	0.438	203	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
4	.753	.039	.019	.134	1.415	0.208	1.035	0.135	0.158	0.539	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
5	.760	.039	.019	.133	1.383	0.199	1.037	0.160	0.160	0.432	196	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI
6	.759	.040	.018	.131	1.439	0.172	1.082	0.138	0.163	0.435	195	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI
7	.753	.038	.020	.138	1.417	0.222	1.036	0.144	0.165	0.593	203	12	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
8	.758	.041	.014	.119	1.434	0.140	1.084	0.097	0.166	0.337	197	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST
9	.752	.039	.017	.129	1.425	0.195	1.046	0.112	0.169	0.552	197	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
10	.761	.039	.020	.139	1.384	0.210	1.036	0.160	0.174	0.468	209	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
11	.753	.039	.020	.135	1.430	0.195	1.044	0.128	0.174	0.575	195	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
12	.757	.038	.022	.142	1.364	0.241	0.992	0.185	0.179	0.613	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
13	.755	.038	.020	.138	1.361	0.233	0.991	0.170	0.179	0.584	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
14	.761	.040	.020	.137	1.402	0.220	1.047	0.161	0.179	0.467	201	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI
15	.761	.040	.020	.138	1.389	0.208	1.050	0.158	0.180	0.478	199	11	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
16	.761	.040	.020	.138	1.377	0.206	1.018	0.158	0.181	0.462	194	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI
17	.756	.038	.021	.142	1.358	0.234	0.988	0.193	0.182	0.563	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
18	.756	.039	.019	.134	1.371	0.250	1.002	0.229	0.183	0.571	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
19	.760	.041	.016	.127	1.396	0.251	1.048	0.202	0.183	0.373	203	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST
20	.759	.041	.016	.126	1.410	0.255	1.055	0.204	0.185	0.376	195	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS
BASE	.719	.044	.000	.033	1.419	0.064	1.041	0.045	0.135	0.048	73	2	ASAR,ASWK

Table A4.

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P	D I F F E R E N C E--							
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.620	.101	.015	.125	1.472	0.204	1.116	0.151	0.454	0.571	102	4	ASAR,ASWK,ASAS,SPRS
2	.617	.107	.028	.158	1.401	0.203	1.017	0.158	0.553	0.634	102	5	ASAR,ASWK,ASAS,CMST,CM1T
3	.617	.107	.024	.145	1.412	0.164	1.029	0.140	0.492	0.708	95	4	ASAR,ASWK,ASAS,CMST
4	.617	.107	.028	.159	1.406	0.204	1.021	0.162	0.542	0.633	102	5	ASAR,ASWK,ASAS,CMST,CM2T
5	.616	.108	.025	.152	1.420	0.208	1.040	0.188	0.496	0.696	101	5	ASAR,ASWK,ASAS,CMST,CMTI
6	.616	.104	.033	.168	1.453	0.441	1.033	0.266	0.551	0.740	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
7	.615	.108	.023	.148	1.380	0.222	1.009	0.178	0.499	0.720	101	5	ASAR,ASWK,ASAS,ASNO,CMST
8	.615	.106	.029	.158	1.386	0.309	0.990	0.177	0.478	0.826	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
9	.615	.104	.033	.167	1.461	0.459	1.039	0.280	0.536	0.768	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
10	.615	.104	.023	.142	1.421	0.207	1.027	0.155	0.518	0.647	100	4	ASAR,ASWK,ASAS,SPOR
11	.614	.105	.029	.154	1.470	0.431	1.049	0.264	0.473	0.842	97	4	ASAR,ASWK,ASAS,ASCS
12	.614	.105	.029	.160	1.485	0.455	1.071	0.294	0.481	0.832	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
13	.611	.102	.026	.150	1.427	0.260	1.028	0.188	0.599	0.656	94	4	ASAR,ASWK,ASAS,CM1T
14	.611	.104	.022	.139	1.431	0.193	1.050	0.174	0.539	0.663	99	4	ASAR,ASWK,ASAS,ASEI
15	.610	.101	.030	.163	1.420	0.242	1.022	0.178	0.604	0.660	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
16	.610	.102	.026	.155	1.441	0.283	1.047	0.218	0.591	0.613	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
17	.610	.106	.006	.101	1.410	0.217	1.106	0.161	0.352	0.302	100	4	ASAR,ASWK,ASEI,SPRS
18	.610	.102	.026	.149	1.437	0.268	1.035	0.198	0.581	0.594	94	4	ASAR,ASWK,ASAS,CM2T
19	.610	.104	.019	.133	1.430	0.216	1.061	0.170	0.519	0.670	101	4	ASAR,ASWK,ASAS,ASGS
20	.610	.102	.027	.158	1.449	0.301	1.053	0.256	0.576	0.592	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A4. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

	Validty Est.		Discr	Brog.	--S U B G R O U P		D I F F	E R E N C E--			Test		
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.616	.104	.033	.168	1.453	0.441	1.033	0.266	0.551	0.740	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
2	.615	.104	.033	.167	1.461	0.459	1.039	0.280	0.536	0.768	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
3	.610	.101	.030	.163	1.420	0.242	1.022	0.178	0.604	0.660	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
4	.614	.105	.029	.160	1.485	0.455	1.071	0.294	0.481	0.832	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
5	.614	.105	.029	.154	1.470	0.431	1.049	0.264	0.473	0.842	97	4	ASAR,ASWK,ASAS,ASCS
6	.615	.106	.029	.158	1.386	0.309	0.990	0.177	0.478	0.826	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
7	.617	.107	.028	.159	1.406	0.204	1.021	0.162	0.542	0.633	102	5	ASAR,ASWK,ASAS,CMST,CM2T
8	.617	.107	.028	.158	1.401	0.203	1.017	0.158	0.553	0.634	102	5	ASAR,ASWK,ASAS,CMST,CM1T
9	.610	.102	.027	.158	1.449	0.301	1.053	0.256	0.576	0.592	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
10	.611	.102	.026	.150	1.427	0.260	1.028	0.188	0.599	0.656	94	4	ASAR,ASWK,ASAS,CM1T
11	.610	.102	.026	.149	1.437	0.268	1.035	0.198	0.581	0.594	94	4	ASAR,ASWK,ASAS,CM2T
12	.610	.102	.026	.155	1.441	0.283	1.047	0.218	0.591	0.613	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
13	.609	.102	.026	.154	1.411	0.368	1.018	0.257	0.601	0.650	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
14	.608	.102	.026	.154	1.421	0.377	1.024	0.267	0.583	0.614	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
15	.616	.108	.025	.152	1.420	0.208	1.040	0.188	0.496	0.696	101	5	ASAR,ASWK,ASAS,CMST,CMTI
16	.617	.107	.024	.145	1.412	0.164	1.029	0.140	0.492	0.708	95	4	ASAR,ASWK,ASAS,CMST
17	.610	.105	.024	.143	1.442	0.263	1.049	0.237	0.501	0.759	103	4	ASAR,ASWK,ASAS,ASPC
18	.604	.109	.024	.150	1.395	0.477	1.026	0.275	0.454	0.521	102	5	ASAR,ASWK,ASCS,ASEI,CM2T
19	.604	.109	.023	.150	1.386	0.481	1.020	0.272	0.467	0.537	102	5	ASAR,ASWK,ASCS,ASEI,CM1T
20	.615	.108	.023	.148	1.380	0.222	1.009	0.178	0.499	0.720	101	5	ASAR,ASWK,ASAS,ASNO,CMST
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A4. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP	DIFFERENCE--							
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.616	.104	.033	.168	1.453	0.441	1.033	0.266	0.551	0.740	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
2	.615	.104	.033	.167	1.461	0.459	1.039	0.280	0.536	0.768	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
3	.610	.101	.030	.163	1.420	0.242	1.022	0.178	0.604	0.660	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
• 4	.614	.105	.029	.160	1.485	0.455	1.071	0.294	0.481	0.832	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
5	.617	.107	.028	.159	1.406	0.204	1.021	0.162	0.542	0.633	102	5	ASAR,ASWK,ASAS,CMST,CM2T
6	.617	.107	.028	.158	1.401	0.203	1.017	0.158	0.553	0.634	102	5	ASAR,ASWK,ASAS,CMST,CM1T
7	.615	.106	.029	.158	1.386	0.309	0.990	0.177	0.478	0.826	103	5	ASAR,ASWK,ASAS,ASCS,ASNO
8	.610	.102	.027	.158	1.449	0.301	1.053	0.256	0.576	0.592	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
9	.610	.102	.026	.155	1.441	0.283	1.047	0.218	0.591	0.613	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
10	.614	.105	.029	.154	1.470	0.431	1.049	0.264	0.473	0.842	97	4	ASAR,ASWK,ASAS,ASCS
11	.609	.102	.026	.154	1.411	0.368	1.018	0.257	0.601	0.650	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
12	.608	.102	.026	.154	1.421	0.377	1.024	0.267	0.583	0.614	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
13	.616	.108	.025	.152	1.420	0.208	1.040	0.188	0.496	0.696	101	5	ASAR,ASWK,ASAS,CMST,CMTI
14	.584	.112	.020	.151	1.358	0.458	0.991	0.262	0.341	0.354	103	6	ASAR,ASWK,ASCS,CMTI,CM1T,CM2T
15	.586	.117	.020	.151	1.335	0.380	0.981	0.229	0.292	0.413	104	6	ASAR,ASWK,ASCS,CMST,CMTI,CM2T
16	.604	.109	.024	.150	1.395	0.477	1.026	0.275	0.454	0.521	102	5	ASAR,ASWK,ASCS,ASEI,CM2T
17	.604	.109	.023	.150	1.386	0.481	1.020	0.272	0.467	0.537	102	5	ASAR,ASWK,ASCS,ASEI,CM1T
18	.611	.102	.026	.150	1.427	0.260	1.028	0.188	0.599	0.656	94	4	ASAR,ASWK,ASAS,CM1T
19	.607	.115	.023	.150	1.354	0.394	1.015	0.236	0.357	0.574	103	5	ASAR,ASWK,ASCS,ASEI,CMST
20	.610	.102	.026	.149	1.437	0.268	1.035	0.198	0.581	0.594	94	4	ASAR,ASWK,ASAS,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A4. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Validty Mean	Est. S.D.	Discr Vldty	Brog. Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Test Time	NV	Predictor Variables
1	.590	.118	.015	.134	1.167	0.363	0.858	0.215	0.212	0.397	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.592	.117	.017	.144	1.188	0.375	0.866	0.221	0.328	0.386	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
3	.584	.118	.011	.124	1.190	0.410	0.895	0.307	0.228	0.403	103	5	ASAR,ASWK,ASNO,ASPC,CMST
4	.586	.114	.013	.121	1.193	0.430	0.868	0.247	0.221	0.376	89	4	ASAR,ASWK,ASCS,ASNO
5	.584	.116	.008	.110	1.195	0.394	0.895	0.278	0.236	0.293	87	4	ASAR,ASWK,ASNO,CMST
6	.592	.117	.018	.145	1.197	0.356	0.872	0.210	0.319	0.408	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
7	.590	.119	.016	.143	1.199	0.346	0.891	0.208	0.237	0.389	103	6	ASAR,ASWK,ASCS,ASNO,CMST,CMTI
8	.587	.115	.011	.123	1.213	0.409	0.899	0.286	0.367	0.358	94	5	ASAR,ASWK,ASNO,CMST,CM1T
9	.590	.112	.016	.135	1.216	0.451	0.878	0.256	0.353	0.371	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
10	.586	.115	.013	.136	1.217	0.395	0.900	0.277	0.379	0.366	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
11	.589	.112	.018	.147	1.218	0.428	0.878	0.243	0.363	0.368	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
12	.587	.115	.011	.125	1.225	0.392	0.906	0.276	0.354	0.396	94	5	ASAR,ASWK,ASNO,CMST,CM2T
13	.589	.112	.015	.134	1.227	0.435	0.885	0.246	0.339	0.375	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
14	.585	.117	.009	.120	1.228	0.388	0.929	0.287	0.263	0.282	93	5	ASAR,ASWK,ASNO,CMST,CMTI
15	.593	.117	.010	.123	1.229	0.421	0.910	0.292	0.268	0.248	100	5	ASAR,ASWK,ASNO,SPOR,CMST
16	.586	.116	.011	.131	1.232	0.402	0.920	0.287	0.368	0.350	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM1T
17	.577	.112	.008	.107	1.232	0.465	0.918	0.358	0.242	0.391	95	4	ASAR,ASWK,ASNO,ASPC
18	.588	.114	.014	.132	1.234	0.431	0.911	0.254	0.253	0.364	95	5	ASAR,ASWK,ASCS,ASNO,CMTI
19	.594	.116	.013	.130	1.235	0.465	0.894	0.270	0.259	0.331	102	5	ASAR,ASWK,ASCS,ASNO,SPOR
20	.589	.113	.016	.142	1.238	0.451	0.904	0.260	0.353	0.352	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM1T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A4. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.590	.118	.015	.134	1.167	0.363	0.858	0.215	0.212	0.397	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.592	.117	.017	.144	1.188	0.375	0.866	0.221	0.328	0.386	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
3	.586	.114	.013	.121	1.193	0.430	0.868	0.247	0.221	0.376	89	4	ASAR,ASWK,ASCS,ASNO
4	.592	.117	.018	.145	1.197	0.356	0.872	0.210	0.319	0.408	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
5	.590	.112	.016	.135	1.216	0.451	0.878	0.256	0.353	0.371	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
6	.589	.112	.018	.147	1.218	0.428	0.878	0.243	0.363	0.368	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
7	.589	.112	.015	.134	1.227	0.435	0.885	0.246	0.339	0.375	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
8	.590	.119	.016	.143	1.199	0.346	0.891	0.208	0.237	0.389	103	6	ASAR,ASWK,ASCS,ASNO,CMST,CMTI
9	.594	.116	.013	.130	1.235	0.465	0.894	0.270	0.259	0.331	102	5	ASAR,ASWK,ASCS,ASNO,SPOR
10	.584	.118	.011	.124	1.190	0.410	0.895	0.307	0.228	0.403	103	5	ASAR,ASWK,ASNO,ASPC,CMST
11	.584	.116	.008	.110	1.195	0.394	0.895	0.278	0.236	0.293	87	4	ASAR,ASWK,ASNO,CMST
12	.587	.115	.011	.123	1.213	0.409	0.899	0.286	0.367	0.358	94	5	ASAR,ASWK,ASNO,CMST,CM1T
13	.586	.115	.013	.136	1.217	0.395	0.900	0.277	0.379	0.366	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
14	.589	.113	.016	.142	1.238	0.451	0.904	0.260	0.353	0.352	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM1T
15	.587	.115	.011	.125	1.225	0.392	0.906	0.276	0.354	0.396	94	5	ASAR,ASWK,ASNO,CMST,CM2T
16	.589	.113	.017	.144	1.247	0.442	0.910	0.281	0.343	0.372	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM2T
17	.593	.117	.010	.123	1.229	0.421	0.910	0.292	0.268	0.248	100	5	ASAR,ASWK,ASNO,SPOR,CMST
18	.588	.114	.014	.132	1.234	0.431	0.911	0.254	0.253	0.364	95	5	ASAR,ASWK,ASCS,ASNO,CMTI
19	.577	.110	.005	.090	1.238	0.474	0.918	0.328	0.255	0.278	79	3	ASAR,ASWK,ASNO
20	.577	.112	.008	.107	1.232	0.465	0.918	0.358	0.242	0.391	95	4	ASAR,ASWK,ASNO,ASPC
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A4. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

	Validty Est.		Discr	Brog.	--S U B G R O U P	D I F F E R E N C E--					Test		
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.596	.113	.007	.107	1.346	0.161	1.048	0.113	0.132	0.171	96	4	ASAR,ASWK,SPRS,CMST
2	.595	.109	.010	.113	1.369	0.373	1.058	0.210	0.135	0.329	98	4	ASAR,ASWK,ASCS,SPRS
3	.583	.118	.017	.132	1.298	0.389	0.957	0.236	0.143	0.441	91	4	ASAR,ASWK,ASCS,CMST
4	.592	.110	.003	.093	1.369	0.228	1.065	0.185	0.145	0.242	104	4	ASAR,ASWK,ASPC,SPRS
5	.580	.120	.011	.118	1.299	0.217	0.962	0.206	0.148	0.368	97	4	ASAR,ASWK,ASPC,CMST
6	.593	.108	.000	.075	1.374	0.211	1.069	0.151	0.149	0.126	88	3	ASAR,ASWK,SPRS
7	.596	.113	.010	.121	1.359	0.217	1.062	0.173	0.149	0.178	102	5	ASAR,ASWK,SPRS,CMST,CMTI
8	.577	.116	.016	.127	1.332	0.466	0.974	0.308	0.149	0.506	99	4	ASAR,ASWK,ASCS,ASPC
9	.605	.114	.009	.115	1.361	0.172	0.944	0.130	0.151	0.211	102	4	ASAR,ASWK,SPAO,CMST
10	.581	.119	.008	.102	1.305	0.187	0.966	0.147	0.152	0.224	81	3	ASAR,ASWK,CMST
11	.578	.114	.014	.118	1.337	0.474	0.975	0.274	0.153	0.419	83	3	ASAR,ASWK,ASCS
12	.596	.109	.013	.128	1.384	0.391	1.075	0.231	0.157	0.298	104	5	ASAR,ASWK,ASCS,SPRS,CMTI
13	.606	.108	.014	.125	1.377	0.447	0.948	0.247	0.157	0.386	104	4	ASAR,ASWK,ASCS,SPAO
14	.602	.110	.002	.084	1.387	0.241	0.957	0.154	0.166	0.141	94	3	ASAR,ASWK,SPAO
15	.594	.108	.004	.099	1.387	0.271	1.084	0.216	0.168	0.152	94	4	ASAR,ASWK,SPRS,CMTI
16	.576	.119	.004	.089	1.297	0.259	0.969	0.171	0.170	0.160	100	3	ASAR,ASWK,ASMK
17	.574	.115	.007	.097	1.332	0.284	0.981	0.242	0.171	0.312	89	3	ASAR,ASWK,ASPC
18	.601	.110	.005	.101	1.396	0.279	0.970	0.235	0.176	0.111	100	4	ASAR,ASWK,SPAO,CMTI
19	.584	.118	.018	.140	1.328	0.378	0.991	0.226	0.177	0.427	97	5	ASAR,ASWK,ASCS,CMST,CMTI
20	.581	.121	.012	.129	1.326	0.245	0.991	0.237	0.177	0.339	103	5	ASAR,ASWK,ASPC,CMST,CMTI
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

Table A5.

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.634	.107	.027	.171	1.446	0.256	1.047	0.161	0.412	0.760	153	8	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPRS,CMST
2	.634	.105	029	.175	1.445	0.273	1.034	0.159	0.406	0.733	163	8	ASAR,ASWK,ASAS,ASCS,ASMC,SPAO,SPRS,CMST
3	.634	.107	.033	.181	1.426	0.311	1.001	0.207	0.436	0.803	160	8	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,SPAO,CMST
4	.634	.105	.033	.177	1.432	0.329	0.996	0.201	0.424	0.811	148	7	ASAR,ASWK,ASAS,ASCS,ASMC,SPAO,CMST
5	.633	.107	.031	.174	1.433	0.310	1.006	0.203	0.428	0.817	138	7	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,CMST
6	.633	.106	.031	.179	1.424	0.339	1.007	0.198	0.423	0.798	162	8	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,SPAO,CMST
7	.633	.106	.028	.178	1.441	0.278	1.041	0.221	0.400	0.751	159	9	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPRS,CMST,CMTI
8	.633	.105	.026	.165	1.454	0.277	1.042	0.157	0.393	0.769	141	7	ASAR,ASWK,ASAS,ASCS,SPAO,SPRS,CMST
9	.633	.106	.025	.166	1.443	0.286	1.054	0.166	0.396	0.759	155	8	ASAR,ASWK,ASAS,ASCS,ASGS,SPAO,SPRS,CMST
10	.633	.107	.028	.177	1.440	0.278	1.041	0.170	0.425	0.652	160	9	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPRS,CMST,CM1T
11	.633	.107	.034	.184	1.428	0.323	0.995	0.236	0.421	0.847	164	8	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,SPAO,CMST
12	.633	.106	.027	.171	1.448	0.254	1.038	0.149	0.393	0.717	154	8	ASAR,ASWK,ASAS,ASCS,SPAO,SPOR,SPRS,CMST
13	.633	.108	.031	.178	1.426	0.296	1.002	0.201	0.429	0.778	151	8	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPOR,CMST
14	.633	.106	.032	.180	1.426	0.315	0.993	0.197	0.425	0.786	161	8	ASAR,ASWK,ASAS,ASCS,ASMC,SPAO,SPOR,CMST
15	.633	.108	.029	.175	1.426	0.321	1.014	0.200	0.426	0.805	152	8	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,SPAO,CMST
16	.633	.109	.029	.171	1.413	0.348	0.993	0.195	0.417	0.823	153	7	ASAR,ASWK,ASAS,ASCS,ASMK,SPAO,CMST
17	.633	.105	.033	.183	1.428	0.338	0.991	0.213	0.414	0.813	154	8	ASAR,ASWK,ASAS,ASCS,ASMC,SPAO,CMST,CMTI
18	.633	.109	.032	.180	1.430	0.306	1.005	0.239	0.425	0.854	154	8	ASAR,ASWK,ASAS,ASCS,ASEI,ASPC,SPAO,CMST
19	.633	.107	.027	.176	1.440	0.292	1.042	0.179	0.421	0.678	160	9	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPRS,CMST,CM2T
20	.633	.104	.024	.161	1.474	0.327	1.062	0.201	0.413	0.751	145	7	ASAR,ASWK,ASAS,ASCS,ASEI,SPAO,SPRS
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A5. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

Rank	Validty Est. Mean	Validty Est. S.D.	Discr Vldty	Brog. Index	--SUBGROUP DIFFERENCE-- Whte-Blck Mean	Whte-Blck Rnge	Whte-Hisp Mean	Whte-Hisp Rnge	Male-Fem. Mean	Male-Fem. Rnge	Test Time	NV	Predictor Variables
1	.621	.107	.041	.204	1.396	0.372	1.007	0.281	0.512	0.784	163	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CMST,CMTI,CM1T,CM2T
2	.617	.104	.041	.198	1.435	0.435	1.030	0.309	0.532	0.770	155	9	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CMTI,CM1T,CM2T
3	.622	.108	.040	.198	1.396	0.361	1.006	0.266	0.518	0.781	157	9	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CMST,CM1T,CM2T
4	.618	.104	.040	.193	1.429	0.427	1.020	0.297	0.539	0.773	149	8	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CM1T,CM2T
5	.621	.106	.040	.193	1.392	0.449	1.009	0.261	0.547	0.734	160	8	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,CM1T,CM2T
6	.619	.106	.040	.196	1.422	0.408	1.025	0.301	0.550	0.766	161	9	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASPC,CM1T,CM2T
7	.618	.105	.040	.197	1.417	0.444	1.031	0.293	0.538	0.748	163	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASPC,CM1T,CM2T
8	.616	.107	.040	.197	1.403	0.478	1.026	0.329	0.544	0.768	160	9	ASAR,ASWK,ASAS,ASCS,ASMK,ASPC,CMTI,CM1T,CM2T
9	.621	.110	.039	.197	1.363	0.394	1.000	0.276	0.528	0.767	162	9	ASAR,ASWK,ASAS,ASCS,ASMK,ASPC,CMST,CM1T,CM2T
10	.616	.106	.039	.200	1.398	0.382	1.004	0.272	0.533	0.766	161	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,ASPC,CMTI,CM1T,CM2T
11	.621	.110	.039	.200	1.360	0.326	0.980	0.231	0.519	0.774	163	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,ASPC,CMST,CM1T,CM2T
12	.619	.110	.039	.200	1.400	0.362	1.022	0.296	0.534	0.776	153	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASPC,CMST,CMTI,CM1T,CM2T
13	.621	.106	.039	.197	1.403	0.363	1.011	0.229	0.518	0.752	147	9	ASAR,ASWK,ASAS,ASCS,ASMC,CMST,CMTI,CM1T,CM2T
14	.618	.108	.039	.201	1.389	0.407	1.024	0.294	0.518	0.756	155	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASPC,CMST,CMTI,CM1T,CM2T
15	.617	.108	.039	.197	1.408	0.377	1.016	0.286	0.521	0.783	141	9	ASAR,ASWK,ASAS,ASCS,ASPC,CMST,CMTI,CM1T,CM2T
16	.618	.103	.039	.192	1.440	0.425	1.032	0.261	0.540	0.728	139	8	ASAR,ASWK,ASAS,ASCS,ASMC,CMTI,CM1T,CM2T
17	.617	.108	.039	.192	1.394	0.464	1.014	0.309	0.551	0.768	154	8	ASAR,ASWK,ASAS,ASCS,ASMK,ASPC,CM1T,CM2T
18	.615	.106	.039	.199	1.426	0.450	1.050	0.327	0.550	0.736	159	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASPC,CMTI,CM1T,CM2T
19	.615	.106	.039	.195	1.439	0.424	1.046	0.326	0.554	0.762	145	9	ASAR,ASWK,ASAS,ASCS,ASEI,ASPC,CMTI,CM1T,CM2T
20	.621	.107	.039	.195	1.404	0.385	1.014	0.290	0.507	0.821	156	9	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CMST,CMTI,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A5. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

Rank	Validty Mean	Est. S.D.	Discr Vldty	Brog. Index	SUBGROUP DIFFERENCE Whte-Blck Mean	Whte-Blck Rnge	Whte-Hisp Mean	Whte-Hisp Rnge	Male-Fem. Mean	Male-Fem. Rnge	Test Time	NV	Predictor Variables
1	.621	.107	.041	.204	1.396	0.372	1.007	0.281	0.512	0.784	163	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASPC,CMST,CMTI,CM1T,CM2T
2	.617	.109	.038	.202	1.340	0.377	0.986	0.232	0.518	0.755	161	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
3	.618	.111	.038	.202	1.353	0.350	0.987	0.261	0.532	0.775	159	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
4	.618	.111	.037	.201	1.346	0.364	0.970	0.252	0.504	0.768	160	11	ASAR,ASWK,ASAS,ASCS,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
5	.618	.108	.039	.201	1.389	0.407	1.024	0.294	0.518	0.756	155	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASPC,CMST,CMTI,CM1T,CM2T
6	.619	.110	.039	.200	1.400	0.362	1.022	0.296	0.534	0.776	153	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASPC,CMST,CMTI,CM1T,CM2T
7	.621	.110	.039	.200	1.360	0.326	0.980	0.231	0.519	0.774	163	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,ASPC,CMST,CM1T,CM2T
8	.619	.109	.039	.200	1.393	0.357	1.004	0.266	0.503	0.778	154	10	ASAR,ASWK,ASAS,ASCS,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
9	.616	.106	.039	.200	1.398	0.382	1.004	0.272	0.533	0.766	161	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,ASPC,CMTI,CM1T,CM2T
10	.621	.108	.039	.200	1.397	0.345	1.017	0.237	0.529	0.745	159	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,CMST,CMTI,CM1T,CM2T
11	.615	.106	.039	.199	1.426	0.450	1.050	0.327	0.550	0.736	159	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASPC,CMTI,CM1T,CM2T
12	.622	.107	.038	.199	1.392	0.352	1.002	0.219	0.505	0.747	160	10	ASAR,ASWK,ASAS,ASCS,ASMC,SPOR,CMST,CMTI,CM1T,CM2T
13	.621	.106	.038	.199	1.390	0.379	1.020	0.227	0.515	0.728	161	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,CMST,CMTI,CM1T,CM2T
14	.621	.109	.036	.199	1.381	0.353	1.019	0.225	0.516	0.714	164	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,SPOR,CMST,CMTI,CM1T,CM2T
15	.617	.110	.038	.199	1.342	0.352	0.969	0.247	0.520	0.777	147	10	ASAR,ASWK,ASAS,ASCS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
16	.616	.109	.036	.199	1.391	0.380	1.007	0.291	0.533	0.757	164	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASNO,ASPC,SPOR,CMTI,CM1T,CM2T
17	.620	.110	.038	.199	1.385	0.374	1.026	0.272	0.535	0.742	161	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASPC,CMST,CM1T,CM2T
18	.617	.106	.038	.199	1.412	0.423	1.031	0.289	0.519	0.746	160	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASPC,SPOR,CMTI,CM1T,CM2T
19	.615	.109	.035	.199	1.359	0.395	1.019	0.313	0.553	0.636	163	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
20	.620	.108	.038	.198	1.367	0.317	0.985	0.194	0.519	0.746	153	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASNO,CMST,CMTI,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A5. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.588	.123	.018	.151	1.155	0.351	0.851	0.240	0.208	0.464	140	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.590	.122	.020	.160	1.176	0.379	0.860	0.249	0.323	0.432	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
3	.590	.122	.022	.170	1.180	0.362	0.863	0.238	0.336	0.428	154	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
4	.591	.121	.021	.163	1.184	0.354	0.864	0.203	0.339	0.406	138	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
5	.590	.123	.021	.161	1.186	0.358	0.867	0.235	0.318	0.453	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
6	.588	.124	.019	.159	1.186	0.356	0.882	0.241	0.231	0.443	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI
7	.587	.122	.016	.147	1.193	0.398	0.893	0.317	0.348	0.427	137	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T
8	.589	.123	.020	.166	1.193	0.378	0.879	0.250	0.324	0.431	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
9	.589	.123	.024	.177	1.194	0.367	0.877	0.245	0.334	0.427	160	10	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
10	.595	.125	.018	.157	1.196	0.392	0.874	0.259	0.244	0.429	153	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST
11	.586	.121	.018	.158	1.197	0.379	0.895	0.304	0.360	0.414	144	8	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
12	.590	.121	.022	.170	1.199	0.349	0.881	0.213	0.340	0.407	144	9	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
13	.590	.121	.019	.158	1.199	0.365	0.883	0.204	0.328	0.412	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T
14	.596	.123	.017	.150	1.200	0.397	0.875	0.228	0.247	0.383	137	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
15	.590	.123	.022	.169	1.200	0.365	0.883	0.245	0.320	0.452	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM2T
16	.588	.118	.019	.153	1.203	0.455	0.874	0.283	0.344	0.432	139	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
17	.587	.122	.017	.149	1.204	0.376	0.900	0.306	0.338	0.419	137	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM2T
18	.591	.121	.021	.162	1.206	0.351	0.886	0.216	0.323	0.424	137	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM2T
19	.585	.123	.015	.146	1.206	0.381	0.920	0.342	0.243	0.382	136	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CMTI
20	.588	.118	.022	.164	1.207	0.437	0.876	0.271	0.356	0.418	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

(Continued)

Table A5. (Cont'd)

On-the-Job Core Technical Proficiency: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P		D I F F E R E N C E--						
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.588	.123	.018	.151	1.155	0.351	0.851	0.240	0.208	0.464	140	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.590	.122	.020	.160	1.176	0.379	0.860	0.249	0.323	0.432	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
3	.590	.122	.022	.170	1.180	0.362	0.863	0.238	0.336	0.428	154	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
4	.591	.121	.021	.163	1.184	0.354	0.864	0.203	0.339	0.406	138	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
5	.590	.123	.021	.161	1.186	0.358	0.867	0.235	0.318	0.453	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
6	.588	.118	.019	.153	1.203	0.455	0.874	0.283	0.344	0.432	139	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
7	.595	.125	.018	.157	1.196	0.392	0.874	0.259	0.244	0.429	153	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST
8	.596	.123	.017	.150	1.200	0.397	0.875	0.228	0.247	0.383	137	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
9	.607	.119	.019	.159	1.258	0.290	0.876	0.205	0.192	0.439	161	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,CMST
10	.588	.118	.022	.164	1.207	0.437	0.876	0.271	0.356	0.418	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T,CM2T
11	.608	.117	.018	.151	1.262	0.294	0.877	0.182	0.194	0.390	145	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST
12	.589	.123	.024	.177	1.194	0.367	0.877	0.245	0.334	0.427	160	10	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
13	.595	.124	.019	.165	1.207	0.407	0.879	0.268	0.322	0.426	160	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST,CM1T
14	.589	.123	.020	.166	1.193	0.378	0.879	0.250	0.324	0.431	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
15	.595	.122	.020	.166	1.212	0.382	0.880	0.216	0.337	0.402	151	9	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CM1T,CM2T
16	.596	.122	.018	.157	1.212	0.403	0.880	0.230	0.326	0.405	144	8	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CM1T
17	.590	.121	.022	.170	1.199	0.349	0.881	0.213	0.340	0.407	144	9	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
18	.606	.118	.019	.164	1.269	0.304	0.881	0.180	0.268	0.381	159	9	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CM1T,CM2T
19	.608	.117	.019	.158	1.270	0.317	0.882	0.184	0.261	0.384	152	8	ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CM1T
20	.588	.124	.019	.159	1.186	0.356	0.882	0.241	0.231	0.443	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A5. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.596	.119	.014	.142	1.312	0.294	1.027	0.196	0.128	0.409	149	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS,CMST
2	.595	.119	.010	.127	1.320	0.209	1.035	0.146	0.129	0.290	139	6	ASAR,ASWK,ASMK,ASPC,SPRS,CMST
3	.594	.116	.011	.130	1.337	0.393	1.043	0.251	0.134	0.405	141	6	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS
4	.607	.117	.012	.140	1.353	0.182	0.988	0.171	0.136	0.287	160	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMST
5	.607	.112	.014	.144	1.367	0.225	0.994	0.214	0.136	0.253	139	7	ASAR,ASWK,ASPC,SPAO,SPRS,CMST,CMTI
6	.608	.116	.011	.131	1.358	0.180	0.991	0.156	0.137	0.170	144	6	ASAR,ASWK,ASMK,SPAO,SPRS,CMST
7	.608	.113	.016	.146	1.359	0.288	0.985	0.184	0.139	0.414	143	7	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMST
8	.608	.112	.018	.156	1.362	0.291	0.989	0.206	0.139	0.395	149	8	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMST,CMTI
9	.609	.115	.016	.146	1.348	0.300	0.982	0.158	0.139	0.359	154	7	ASAR,ASWK,ASCS,ASMK,SPAO,SPRS,CMST
10	.608	.116	.013	.142	1.361	0.232	0.995	0.219	0.140	0.150	150	7	ASAR,ASWK,ASMK,SPAO,SPRS,CMST,CMTI
11	.582	.124	.018	.147	1.258	0.385	0.943	0.267	0.141	0.511	134	6	ASAR,ASWK,ASCS,ASMK,ASPC,CMST
12	.609	.115	.018	.156	1.354	0.302	0.988	0.216	0.143	0.337	160	8	ASAR,ASWK,ASCS,ASMK,SPAO,SPRS,CMST,CMTI
13	.607	.113	.014	.141	1.364	0.402	0.991	0.246	0.143	0.404	162	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPRS
14	.595	.119	.013	.142	1.331	0.251	1.046	0.193	0.143	0.254	145	7	ASAR,ASWK,ASMK,ASPC,SPRS,CMST,CMTI
15	.607	.109	.012	.132	1.383	0.378	0.999	0.238	0.143	0.404	135	6	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS
16	.596	.119	.017	.154	1.327	0.308	1.041	0.212	0.144	0.392	155	8	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS,CMST,CMTI
17	.608	.112	.013	.134	1.367	0.398	0.993	0.207	0.144	0.334	146	6	ASAR,ASWK,ASCS,ASMK,SPAO,SPRS
18	.606	.114	.008	.123	1.370	0.252	1.000	0.213	0.144	0.235	152	6	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS
19	.607	.108	.015	.145	1.387	0.404	1.005	0.271	0.146	0.369	141	7	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMTI
20	.606	.113	.012	.141	1.370	0.326	1.002	0.275	0.146	0.192	158	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMTI
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

Table A6.

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty Est.		Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.634	.110	.029	.182	1.416	0.292	1.032	0.181	0.424	0.723	202	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPRS,CMST
2	.633	.109	.028	.180	1.417	0.316	1.037	0.175	0.409	0.706	204	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,SPAO,SPRS,CMST
3	.633	.111	.034	.191	1.397	0.344	0.994	0.249	0.439	0.838	203	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,SPAO,CMST
4	.633	.112	.028	.181	1.418	0.295	1.036	0.210	0.417	0.790	196	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMK,ASPC,SPAO,SPRS,CMST
5	.633	.108	.030	.185	1.416	0.320	1.021	0.231	0.400	0.725	196	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,SPAO,SPRS,CMST,CMTI
6	.633	.109	.030	.189	1.412	0.306	1.027	0.242	0.412	0.721	208	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPRS,CMST,CMTI
7	.633	.110	.030	.184	1.416	0.310	1.024	0.196	0.411	0.764	206	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPAO,SPRS,CMST
8	.633	.112	.030	.188	1.412	0.296	1.029	0.199	0.424	0.760	218	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,SPAO,SPRS,CMST
9	.633	.109	.029	.182	1.418	0.301	1.025	0.164	0.410	0.694	203	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,SPAO,SPOR,SPRS,CMST
10	.633	.107	.027	.175	1.439	0.365	1.044	0.220	0.425	0.716	194	9	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPRS
11	.633	.110	.029	.186	1.414	0.286	1.030	0.175	0.422	0.688	215	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPOR,SPRS,CMST
12	.633	.110	.025	.176	1.419	0.297	1.046	0.171	0.415	0.738	194	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,SPAO,SPRS,CMST
13	.633	.110	.028	.184	1.413	0.303	1.038	0.181	0.419	0.701	216	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,SPAO,SPRS,CMST
14	.633	.111	.032	.188	1.399	0.338	0.993	0.211	0.440	0.771	200	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPOR,CMST
15	.633	.110	.031	.186	1.399	0.348	1.001	0.207	0.437	0.779	201	10	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,SPAO,CMST
16	.632	.106	.029	.186	1.425	0.295	1.038	0.230	0.402	0.703	195	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,SPAO,SPRS,CMST,CMTI
17	.632	.109	.029	.187	1.415	0.303	1.031	0.227	0.402	0.693	199	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMK,SPAO,SPOR,SPRS,CMST,CMTI
18	.632	.109	.029	.183	1.415	0.328	1.022	0.169	0.421	0.648	197	10	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,SPAO,SPRS,CMST,CMIT
19	.632	.110	.029	.187	1.411	0.315	1.027	0.179	0.431	0.643	209	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,SPAO,SPRS,CMST,CMIT
20	.632	.110	.030	.189	1.411	0.342	1.029	0.259	0.404	0.778	202	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMK,ASPC,SPAO,SPRS,CMST,CMTI
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A6. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

	Validty Est.		Discr	Brog.	--S U B G R O U P D I F F E R E N C E--						Test		
					Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.622	.111	.042	.211	1.357	0.405	1.002	0.305	0.525	0.766	202	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
2	.621	.110	.042	.211	1.353	0.439	1.003	0.298	0.513	0.749	204	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
3	.619	.109	.042	.207	1.390	0.460	1.023	0.329	0.543	0.761	194	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,CMTI,CM1T,CM2T
4	.622	.111	.041	.211	1.357	0.414	0.990	0.283	0.502	0.764	203	12	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
5	.618	.108	.041	.207	1.388	0.487	1.024	0.319	0.532	0.748	196	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,CMTI,CM1T,CM2T
6	.621	.111	.041	.213	1.351	0.432	1.006	0.307	0.521	0.739	216	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
7	.618	.109	.041	.209	1.384	0.479	1.026	0.327	0.540	0.737	208	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMTI,CM1T,CM2T
8	.623	.112	.041	.206	1.356	0.388	1.001	0.283	0.531	0.764	196	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,CMST,CM1T,CM2T
9	.620	.109	.041	.206	1.389	0.458	1.010	0.299	0.519	0.763	195	11	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPOR,CMTI,CM1T,CM2T
10	.622	.113	.041	.213	1.354	0.404	0.996	0.295	0.512	0.756	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
11	.620	.111	.041	.209	1.356	0.381	0.993	0.242	0.516	0.772	196	12	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
12	.623	.110	.041	.206	1.355	0.419	1.006	0.275	0.519	0.745	198	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,CMST,CM1T,CM2T
13	.621	.111	.041	.213	1.349	0.431	0.997	0.286	0.500	0.745	217	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
14	.619	.109	.040	.205	1.379	0.465	1.018	0.307	0.547	0.737	202	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CM1T,CM2T
15	.620	.110	.040	.208	1.385	0.444	1.015	0.309	0.528	0.755	207	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,SPOR,CMTI,CM1T,CM2T
16	.623	.111	.040	.205	1.358	0.402	0.992	0.270	0.509	0.759	197	11	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPOR,CMST,CM1T,CM2T
17	.622	.112	.040	.208	1.352	0.408	1.007	0.278	0.527	0.736	210	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CM1T,CM2T
18	.619	.109	.040	.208	1.381	0.464	1.015	0.298	0.517	0.746	209	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,SPOR,CMTI,CM1T,CM2T
19	.620	.111	.040	.209	1.353	0.430	1.013	0.315	0.531	0.739	194	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
20	.619	.111	.040	.212	1.350	0.407	1.001	0.237	0.513	0.747	210	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A6. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE-- Whte-Blck		Whte-Hisp		Male-Fem.				
Rank	Validty Mean	Est. S.D.	Discr Vldty	Brog. Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Test Time	NV	Predictor Variables
1	.618	.113	.039	.214	1.352	0.407	1.008	0.259	0.522	0.741	222	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
2	.621	.111	.041	.213	1.351	0.432	1.006	0.307	0.521	0.739	216	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
3	.619	.112	.039	.213	1.348	0.411	0.997	0.236	0.500	0.744	223	14	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
4	.620	.114	.039	.213	1.356	0.389	0.999	0.258	0.512	0.758	221	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
5	.622	.113	.041	.213	1.354	0.404	0.996	0.295	0.512	0.756	215	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
6	.621	.111	.041	.213	1.349	0.431	0.997	0.286	0.500	0.745	217	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
7	.619	.112	.038	.212	1.356	0.390	0.997	0.266	0.506	0.741	208	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
8	.619	.111	.040	.212	1.350	0.407	1.001	0.237	0.513	0.747	210	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
9	.620	.112	.040	.211	1.354	0.390	0.988	0.243	0.502	0.763	209	13	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
10	.620	.113	.040	.211	1.358	0.380	1.004	0.259	0.525	0.768	208	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
11	.621	.110	.042	.211	1.353	0.439	1.003	0.298	0.513	0.749	204	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
12	.622	.111	.042	.211	1.357	0.405	1.002	0.305	0.525	0.766	202	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
13	.628	.108	.036	.211	1.389	0.356	0.986	0.268	0.432	0.697	223	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
14	.618	.110	.040	.211	1.379	0.458	1.018	0.307	0.525	0.734	221	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASMK,ASPC,SPOR,CMTI,CM1T,CM2T
15	.621	.110	.039	.211	1.370	0.380	1.008	0.280	0.506	0.743	202	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
16	.622	.111	.041	.211	1.357	0.414	0.990	0.283	0.502	0.764	203	12	ASAR,ASWK,ASAS,ASCS,ASMC,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
17	.620	.112	.039	.211	1.350	0.423	1.006	0.299	0.514	0.734	207	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
18	.618	.113	.037	.210	1.347	0.411	1.003	0.254	0.515	0.736	213	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
19	.618	.111	.039	.210	1.359	0.385	1.001	0.265	0.520	0.747	195	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
20	.627	.109	.036	.210	1.389	0.368	0.984	0.242	0.436	0.704	216	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMC,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A6. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.604	.121	.021	.182	1.272	0.338	0.887	0.229	0.262	0.399	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
2	.605	.116	.022	.185	1.280	0.337	0.994	0.185	0.304	0.410	202	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
3	.605	.117	.019	.177	1.282	0.318	0.998	0.132	0.311	0.412	196	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CM1T,CM2T
4	.606	.117	.019	.175	1.285	0.340	0.999	0.161	0.296	0.416	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
5	.605	.118	.019	.177	1.285	0.285	0.926	0.186	0.236	0.357	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
6	.605	.117	.019	.177	1.288	0.337	1.000	0.191	0.297	0.342	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM2T
7	.609	.115	.024	.186	1.288	0.349	0.923	0.210	0.287	0.443	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
8	.605	.120	.018	.180	1.289	0.295	0.928	0.191	0.237	0.362	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
9	.605	.120	.016	.172	1.290	0.290	0.929	0.168	0.244	0.353	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
10	.609	.116	.023	.189	1.293	0.359	0.928	0.221	0.288	0.450	208	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
11	.606	.120	.017	.173	1.293	0.299	0.931	0.178	0.231	0.371	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
12	.607	.120	.016	.166	1.293	0.295	0.931	0.171	0.238	0.361	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
13	.608	.119	.018	.169	1.293	0.284	0.935	0.168	0.190	0.343	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
14	.606	.120	.017	.174	1.295	0.297	0.933	0.196	0.227	0.336	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
15	.606	.120	.015	.165	1.295	0.291	0.933	0.167	0.233	0.331	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
16	.611	.116	.022	.182	1.297	0.365	0.930	0.223	0.282	0.455	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
17	.606	.119	.026	.193	1.297	0.429	0.982	0.250	0.428	0.547	199	13	ASAR,ASWK,ASCS,ASEI,ASGS,ASMK,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
18	.604	.119	.016	.173	1.298	0.331	0.945	0.231	0.243	0.352	199	12	ASAR,ASWK,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
19	.610	.117	.022	.182	1.298	0.344	0.935	0.203	0.295	0.445	202	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
20	.608	.116	.020	.180	1.298	0.388	0.940	0.290	0.295	0.456	198	12	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A6. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.604	.121	.021	.182	1.272	0.338	0.887	0.229	0.262	0.399	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
2	.609	.115	024	.186	1.288	0.349	0.923	0.210	0.287	0.443	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
3	.605	.118	.019	.177	1.285	0.285	0.926	0.186	0.236	0.357	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
4	.609	.116	.023	.189	1.293	0.359	0.928	0.221	0.288	0.450	208	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
5	.605	.120	.018	.180	1.289	0.295	0.928	0.191	0.237	0.362	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
6	.605	.120	.016	.172	1.290	0.290	0.929	0.168	0.244	0.353	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
7	.611	.116	.022	.182	1.297	0.365	0.930	0.223	0.282	0.455	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
8	.606	.120	.017	.173	1.293	0.299	0.931	0.178	0.231	0.371	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
9	.607	.120	.016	.166	1.293	0.295	0.931	0.171	0.238	0.361	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
10	.610	.117	.021	.180	1.300	0.357	0.933	0.222	0.282	0.453	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM2T
11	.606	.120	.017	.174	1.295	0.297	0.933	0.196	0.227	0.336	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
12	.606	.120	.015	.165	1.295	0.291	0.933	0.167	0.233	0.331	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
13	.608	.119	.018	.169	1.293	0.284	0.935	0.168	0.190	0.343	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
14	.610	.117	.022	.182	1.298	0.344	0.935	0.203	0.295	0.445	202	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
15	.611	.117	.020	.174	1.301	0.340	0.935	0.216	0.249	0.477	194	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI
16	.612	.117	.021	.176	1.302	0.352	0.937	0.208	0.290	0.449	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T
17	.613	.119	.027	.195	1.334	0.358	0.937	0.235	0.320	0.519	216	13	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
18	.614	.118	.028	.193	1.336	0.356	0.937	0.235	0.322	0.523	203	12	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
19	.611	.117	.019	.173	1.305	0.341	0.939	0.202	0.288	0.447	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM2T
20	.608	.116	.020	.180	1.298	0.388	0.940	0.290	0.295	0.456	198	12	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A6. (Cont'd)

On-the-Job Core Technical Proficiency Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.608	.119	.018	.169	1.293	0.284	0.935	0.168	0.190	0.343	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
2	.606	.119	.019	.173	1.344	0.311	0.973	0.245	0.204	0.326	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
3	.607	.118	.019	.172	1.340	0.308	0.970	0.223	0.209	0.338	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
4	.614	.114	.020	.172	1.346	0.349	1.004	0.215	0.212	0.397	203	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
5	.605	.119	.020	.179	1.337	0.304	0.968	0.237	0.216	0.326	203	12	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
6	.613	.114	.018	.169	1.314	0.292	0.976	0.175	0.216	0.415	196	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI
7	.614	.111	.018	.165	1.370	0.434	1.018	0.258	0.219	0.386	195	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI
8	.605	.115	.018	.172	1.354	0.404	0.979	0.283	0.223	0.324	195	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
9	.606	.119	.018	.170	1.339	0.303	0.970	0.185	0.223	0.317	197	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
10	.615	.115	.018	.163	1.352	0.315	1.011	0.172	0.224	0.415	197	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST
11	.613	.116	.018	.173	1.318	0.301	0.977	0.176	0.227	0.395	209	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
12	.606	.120	.017	.174	1.295	0.297	0.933	0.196	0.227	0.336	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
13	.612	.115	.017	.168	1.323	0.321	0.992	0.202	0.229	0.336	199	11	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
14	.606	.120	.017	.173	1.293	0.299	0.931	0.178	0.231	0.371	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
15	.612	.112	.016	.166	1.343	0.388	0.993	0.220	0.232	0.388	201	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI
16	.606	.120	.015	.165	1.295	0.291	0.933	0.167	0.233	0.331	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
17	.606	.116	.015	.166	1.318	0.400	0.948	0.243	0.233	0.318	194	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI,CM2T
18	.605	.118	.019	.177	1.285	0.285	0.926	0.186	0.236	0.357	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
19	.605	.120	.018	.180	1.289	0.295	0.928	0.191	0.237	0.362	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
20	.613	.113	.020	.173	1.348	0.381	1.005	0.250	0.238	0.385	197	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM2T
BASE	.574	.113	.003	.071	1.339	0.247	0.984	0.174	0.180	0.156	73	2	ASAR,ASWK

(Continued)

Table A7.

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P	D I F F E R E N C E--							
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.491	.101	.021	.134	1.361	0.381	1.043	0.256	0.634	0.632	102	4	ASAR,ASWK,ASAS,SPRS
2	.489	.098	.035	.161	1.330	0.454	0.971	0.387	0.858	0.921	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
3	.489	.099	.038	.166	1.341	0.492	0.977	0.456	0.836	0.795	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
4	.488	.107	.029	.153	1.333	0.348	0.988	0.286	0.714	0.744	101	5	ASAR,ASWK,ASAS,CMST,CMTI
5	.487	.105	.034	.158	1.283	0.425	0.918	0.319	0.817	0.875	102	5	ASAR,ASWK,ASAS,CMST,CM2T
6	.487	.104	.031	.154	1.274	0.518	0.914	0.385	0.841	0.977	102	5	ASAR,ASWK,ASAS,CMST,CM1T
7	.486	.104	.028	.150	1.255	0.518	0.942	0.358	0.770	0.732	99	5	ASAR,ASWK,ASAS,ASNO,CMTI
8	.486	.109	.027	.149	1.164	0.597	0.857	0.410	0.725	0.744	101	5	ASAR,ASWK,ASAS,ASNO,CMST
9	.486	.113	.023	.137	1.290	0.402	0.923	0.291	0.735	0.819	100	4	ASAR,ASWK,ASAS,SPOR
10	.485	.099	.031	.154	1.194	0.680	0.860	0.481	0.907	1.002	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
11	.484	.104	.027	.143	1.378	0.325	1.020	0.297	0.745	0.740	93	4	ASAR,ASWK,ASAS,CMTI
12	.484	.100	.033	.156	1.207	0.584	0.867	0.404	0.877	0.878	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
13	.483	.101	.030	.148	1.304	0.588	0.927	0.435	0.892	1.064	94	4	ASAR,ASWK,ASAS,CM1T
14	.482	.107	.028	.149	1.387	0.345	1.020	0.291	0.725	0.840	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
15	.482	.101	.034	.160	1.295	0.561	0.919	0.420	0.896	1.062	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
16	.482	.102	.032	.148	1.320	0.478	0.937	0.359	0.862	0.920	94	4	ASAR,ASWK,ASAS,CM2T
17	.482	.110	.025	.139	1.302	0.443	0.943	0.319	0.701	0.754	95	4	ASAR,ASWK,ASAS,CMST
18	.482	.097	.024	.138	1.280	0.391	0.953	0.321	0.603	0.564	101	4	ASAR,ASWK,ASMC,CMTI
19	.481	.103	.030	.152	1.317	0.617	0.930	0.446	0.870	1.143	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
20	.481	.104	.026	.141	1.308	0.592	0.984	0.355	0.730	0.737	101	4	ASAR,ASWK,ASAS,ASGS
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A7. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.489	.099	.038	.166	1.341	0.492	0.977	0.456	0.836	0.795	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
2	.489	.098	.035	.161	1.330	0.454	0.971	0.387	0.858	0.921	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
3	.482	.101	.034	.160	1.295	0.561	0.919	0.420	0.896	1.062	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
4	.487	.105	.034	.158	1.283	0.425	0.918	0.319	0.817	0.875	102	5	ASAR,ASWK,ASAS,CMST,CM2T
5	.484	.100	.033	.156	1.207	0.584	0.867	0.404	0.877	0.878	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
6	.482	.102	.032	.148	1.320	0.478	0.937	0.359	0.862	0.920	94	4	ASAR,ASWK,ASAS,CM2T
7	.487	.104	.031	.154	1.274	0.518	0.914	0.385	0.841	0.977	102	5	ASAR,ASWK,ASAS,CMST,CM1T
8	.480	.105	.031	.151	1.332	0.520	0.940	0.373	0.842	0.996	104	5	ASAR,ASWK,ASAS,ASCS,CM2T
9	.485	.099	.031	.154	1.194	0.680	0.860	0.481	0.907	1.002	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
10	.483	.101	.030	.148	1.304	0.588	0.927	0.435	0.892	1.064	94	4	ASAR,ASWK,ASAS,CM1T
11	.481	.103	.030	.152	1.317	0.617	0.930	0.446	0.870	1.143	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
12	.488	.107	.029	.153	1.333	0.348	0.988	0.286	0.714	0.744	101	5	ASAR,ASWK,ASAS,CMST,CMTI
13	.461	.108	.029	.161	1.028	0.667	0.785	0.536	0.596	0.491	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM2T
14	.482	.107	.028	.149	1.387	0.345	1.020	0.291	0.725	0.840	103	5	ASAR,ASWK,ASAS,ASCS,CMTI
15	.486	.104	.028	.150	1.255	0.518	0.942	0.358	0.770	0.732	99	5	ASAR,ASWK,ASAS,ASNO,CMTI
16	.475	.097	.028	.158	1.174	0.593	0.904	0.506	0.762	0.679	104	6	ASAR,ASWK,ASEI,ASNO,CMTI,CM2T
17	.484	.104	.027	.143	1.378	0.325	1.020	0.297	0.745	0.740	93	4	ASAR,ASWK,ASAS,CMTI
18	.486	.109	.027	.149	1.164	0.597	0.857	0.410	0.725	0.744	101	5	ASAR,ASWK,ASAS,ASNO,CMST
19	.473	.097	.027	.150	1.279	0.519	0.977	0.462	0.730	0.654	98	5	ASAR,ASWK,ASEI,CMTI,CM2T
20	.444	.112	.026	.157	1.207	0.579	0.907	0.507	0.531	0.520	104	6	ASAR,ASWK,ASCS,CMST,CMTI,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A7. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.489	.099	.038	.166	1.341	0.492	0.977	0.456	0.836	0.795	100	5	ASAR,ASWK,ASAS,CMTI,CM2T
2	.489	.098	.035	.161	1.330	0.454	0.971	0.387	0.858	0.921	100	5	ASAR,ASWK,ASAS,CMTI,CM1T
3	.461	.108	.029	.161	1.028	0.667	0.785	0.536	0.596	0.491	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM2T
4	.482	.101	.034	.160	1.295	0.561	0.919	0.420	0.896	1.062	101	5	ASAR,ASWK,ASAS,CM1T,CM2T
5	.487	.105	.034	.158	1.283	0.425	0.918	0.319	0.817	0.875	102	5	ASAR,ASWK,ASAS,CMST,CM2T
6	.475	.097	.028	.158	1.174	0.593	0.904	0.506	0.762	0.679	104	6	ASAR,ASWK,ASEI,ASNO,CMTI,CM2T
7	.444	.111	.026	.157	1.221	0.562	0.904	0.517	0.536	0.552	101	6	ASAR,ASWK,CMST,CMTI,CM1T,CM2T
8	.444	.112	.026	.157	1.207	0.579	0.907	0.507	0.531	0.520	104	6	ASAR,ASWK,ASCS,CMST,CMTI,CM2T
9	.484	.100	.033	.156	1.207	0.584	0.867	0.404	0.877	0.878	100	5	ASAR,ASWK,ASAS,ASNO,CM2T
10	.458	.103	.025	.155	1.058	0.708	0.802	0.562	0.657	0.618	99	6	ASAR,ASWK,ASNO,CMTI,CM1T,CM2T
11	.443	.109	.025	.155	1.222	0.599	0.911	0.515	0.590	0.682	103	6	ASAR,ASWK,ASCS,CMTI,CM1T,CM2T
12	.460	.111	.025	.155	1.013	0.619	0.777	0.467	0.609	0.623	100	6	ASAR,ASWK,ASNO,CMST,CMTI,CM1T
13	.461	.108	.025	.155	1.040	0.637	0.775	0.519	0.554	0.507	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM2T
14	.487	.104	.031	.154	1.274	0.518	0.914	0.385	0.841	0.977	102	5	ASAR,ASWK,ASAS,CMST,CM1T
15	.485	.099	.031	.154	1.194	0.680	0.860	0.481	0.907	1.002	100	5	ASAR,ASWK,ASAS,ASNO,CM1T
16	.443	.114	.024	.154	1.186	0.542	0.896	0.448	0.550	0.687	104	6	ASAR,ASWK,ASCS,CMST,CMTI,CM1T
17	.462	.111	.024	.153	1.022	0.580	0.764	0.481	0.565	0.620	102	6	ASAR,ASWK,ASCS,ASNO,CMTI,CM1T
18	.488	.107	.029	.153	1.333	0.348	0.988	0.286	0.714	0.744	101	5	ASAR,ASWK,ASAS,CMST,CMTI
19	.475	.099	.024	.153	1.161	0.546	0.897	0.449	0.780	0.791	104	6	ASAR,ASWK,ASEI,ASNO,CMTI,CM1T
20	.481	.103	.030	.152	1.317	0.617	0.930	0.446	0.870	1.143	104	5	ASAR,ASWK,ASAS,ASCS,CM1T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A7. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.443	.122	.012	.126	0.916	0.711	0.683	0.482	0.311	0.268	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.437	.119	.013	.128	0.936	0.753	0.717	0.514	0.353	0.267	103	5	ASAR,ASWK,ASNO,ASPC,CMST
3	.440	.118	.013	.120	0.940	0.759	0.720	0.529	0.355	0.270	87	4	ASAR,ASWK,ASNO,CMST
4	.437	.119	.009	.113	0.947	0.710	0.695	0.486	0.322	0.289	89	4	ASAR,ASWK,ASCS,ASNO
5	.455	.116	.017	.141	0.954	0.690	0.697	0.498	0.541	0.616	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
6	.453	.112	.017	.135	0.972	0.741	0.726	0.541	0.605	0.642	94	5	ASAR,ASWK,ASNO,CMST,CM1T
7	.455	.114	.018	.141	0.974	0.629	0.709	0.448	0.528	0.497	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
8	.455	.120	.011	.127	0.982	0.751	0.737	0.517	0.406	0.251	100	5	ASAR,ASWK,ASNO,SPOR,CMST
9	.453	.112	.016	.132	0.983	0.700	0.711	0.506	0.563	0.645	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
10	.428	.114	.008	.111	0.985	0.778	0.746	0.540	0.373	0.292	95	4	ASAR,ASWK,ASNO,ASPC
11	.455	.120	.020	.148	0.986	0.535	0.758	0.416	0.382	0.335	103	6	ASAR,ASWK,ASCS,ASNO,CMST,CMTI
12	.453	.110	.020	.145	0.986	0.745	0.732	0.541	0.626	0.631	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
13	.431	.113	.009	.103	0.988	0.778	0.746	0.544	0.382	0.311	79	3	ASAR,ASWK,ASNO
14	.454	.109	.020	.138	0.993	0.678	0.739	0.488	0.587	0.564	94	5	ASAR,ASWK,ASNO,CMST,CM2T
15	.452	.123	.008	.121	0.994	0.702	0.724	0.477	0.383	0.271	102	5	ASAR,ASWK,ASCS,ASNO,SPOR
16	.452	.110	.017	.141	0.996	0.706	0.716	0.509	0.584	0.637	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
17	.452	.110	.015	.128	1.006	0.629	0.724	0.450	0.546	0.526	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
18	.458	.111	.014	.129	1.009	0.722	0.822	0.447	0.436	0.363	101	5	ASAR,ASWK,ASGS,ASNO,CMST
19	.447	.107	.015	.132	1.010	0.769	0.747	0.554	0.642	0.649	102	5	ASAR,ASWK,ASNO,ASPC,CM1T
20	.452	.116	.019	.141	1.010	0.577	0.795	0.375	0.426	0.337	93	5	ASAR,ASWK,ASNO,CMST,CMTI
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A7. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.443	.122	.012	.126	0.916	0.711	0.683	0.482	0.311	0.268	97	5	ASAR,ASWK,ASCS,ASNO,CMST
2	.437	.119	.009	.113	0.947	0.710	0.695	0.486	0.322	0.289	89	4	ASAR,ASWK,ASCS,ASNO
3	.455	.116	.017	.141	0.954	0.690	0.697	0.498	0.541	0.616	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM1T
4	.455	.114	.018	.141	0.974	0.629	0.709	0.448	0.528	0.497	104	6	ASAR,ASWK,ASCS,ASNO,CMST,CM2T
5	.453	.112	.016	.132	0.983	0.700	0.711	0.506	0.563	0.645	96	5	ASAR,ASWK,ASCS,ASNO,CM1T
6	.452	.110	.017	.141	0.996	0.706	0.716	0.509	0.584	0.637	103	6	ASAR,ASWK,ASCS,ASNO,CM1T,CM2T
7	.437	.119	.013	.128	0.936	0.753	0.717	0.514	0.353	0.267	103	5	ASAR,ASWK,ASNO,ASPC,CMST
8	.440	.118	.013	.120	0.940	0.759	0.720	0.529	0.355	0.270	87	4	ASAR,ASWK,ASNO,CMST
9	.452	.123	.008	.121	0.994	0.702	0.724	0.477	0.383	0.271	102	5	ASAR,ASWK,ASCS,ASNO,SPOR
10	.452	.110	.015	.128	1.006	0.629	0.724	0.450	0.546	0.526	96	5	ASAR,ASWK,ASCS,ASNO,CM2T
11	.453	.112	.017	.135	0.972	0.741	0.726	0.541	0.605	0.642	94	5	ASAR,ASWK,ASNO,CMST,CM1T
12	.453	.110	.020	.145	0.986	0.745	0.732	0.541	0.626	0.631	101	6	ASAR,ASWK,ASNO,CMST,CM1T,CM2T
13	.455	.120	.011	.127	0.982	0.751	0.737	0.517	0.406	0.251	100	5	ASAR,ASWK,ASNO,SPOR,CMST
14	.454	.109	.020	.138	0.993	0.678	0.739	0.488	0.587	0.564	94	5	ASAR,ASWK,ASNO,CMST,CM2T
15	.428	.114	.008	.111	0.985	0.778	0.746	0.540	0.373	0.292	95	4	ASAR,ASWK,ASNO,ASPC
16	.431	.113	.009	.103	0.988	0.778	0.746	0.544	0.382	0.311	79	3	ASAR,ASWK,ASNO
17	.468	.107	.011	.117	1.091	0.690	0.747	0.495	0.320	0.281	100	4	ASAR,ASWK,ASNO,SPAO
18	.447	.107	.015	.132	1.010	0.769	0.747	0.554	0.642	0.649	102	5	ASAR,ASWK,ASNO,ASPC,CM1T
19	.449	.106	.014	.122	1.015	0.769	0.749	0.561	0.647	0.683	86	4	ASAR,ASWK,ASNO,CM1T
20	.449	.104	.016	.134	1.027	0.775	0.754	0.562	0.667	0.676	93	5	ASAR,ASWK,ASNO,CM1T,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A7. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 74-104 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.444	.110	.014	.123	1.225	0.480	0.983	0.313	0.145	0.168	96	4	ASAR,ASWK,SPRS,CMST
2	.442	.107	.009	.103	1.254	0.433	1.005	0.270	0.164	0.124	88	3	ASAR,ASWK,SPRS
3	.458	.115	.013	.122	1.238	0.439	0.830	0.351	0.171	0.161	102	4	ASAR,ASWK,SPAO,CMST
4	.439	.108	.007	.111	1.247	0.432	0.997	0.311	0.173	0.244	104	4	ASAR,ASWK,ASPC,SPRS
5	.417	.124	.009	.105	1.202	0.536	0.893	0.377	0.179	0.211	81	3	ASAR,ASWK,CMST
6	.456	.112	.007	.100	1.265	0.389	0.844	0.330	0.187	0.139	94	3	ASAR,ASWK,SPAO
7	.420	.131	.005	.098	1.141	0.607	0.876	0.388	0.189	0.178	100	3	ASAR,ASWK,ASMK
8	.413	.126	.008	.113	1.191	0.539	0.883	0.387	0.190	0.341	97	4	ASAR,ASWK,ASPC,CMST
9	.441	.109	.011	.116	1.229	0.538	0.995	0.296	0.195	0.256	98	4	ASAR,ASWK,ASCS,SPRS
10	.416	.125	.010	.116	1.166	0.640	0.878	0.402	0.217	0.303	91	4	ASAR,ASWK,ASCS,CMST
11	.455	.105	.025	.149	1.258	0.485	1.022	0.371	0.220	0.214	102	5	ASAR,ASWK,SPRS,CMST,CMTI
12	.407	.125	.002	.092	1.238	0.506	0.910	0.377	0.224	0.292	89	3	ASAR,ASWK,ASPC
13	.456	.113	.010	.115	1.235	0.488	0.831	0.353	0.226	0.286	104	4	ASAR,ASWK,ASCS,SPAO
14	.455	.101	.021	.137	1.281	0.491	1.040	0.376	0.231	0.250	94	4	ASAR,ASWK,SPRS,CMTI
15	.409	.126	.005	.097	1.226	0.589	0.907	0.395	0.237	0.370	83	3	ASAR,ASWK,ASCS
16	.405	.127	.004	.106	1.219	0.576	0.901	0.399	0.238	0.353	99	4	ASAR,ASWK,ASCS,ASPC
17	.465	.108	.018	.133	1.289	0.365	0.894	0.362	0.240	0.258	100	4	ASAR,ASWK,SPAO,CMTI
18	.455	.114	.009	.115	1.208	0.387	0.940	0.228	0.253	0.137	101	4	ASAR,ASWK,SPOR,SPRS
19	.454	.103	.022	.144	1.257	0.497	1.031	0.369	0.266	0.364	104	5	ASAR,ASWK,ASCS,SPRS,CMTI
20	.432	.120	.017	.130	1.255	0.395	0.963	0.365	0.274	0.211	87	4	ASAR,ASWK,CMST,CMTI
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

Table A8.

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P D I F F E R E N C E--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.515	.101	.038	.185	1.261	0.525	0.931	0.321	0.652	0.810	163	8	ASAR,ASWK,ASAS,ASMK,SPAO,SPRS,CMTI,CM1T
2	.514	.101	.040	.187	1.237	0.635	0.899	0.358	0.687	0.846	162	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMTI,CM1T
3	.514	.102	.039	.186	1.268	0.525	0.935	0.373	0.645	0.667	163	8	ASAR,ASWK,ASAS,ASMK,SPAO,SPRS,CMTI,CM2T
4	.514	.105	.032	.171	1.265	0.435	0.907	0.323	0.641	0.740	163	7	ASAR,ASWK,ASAS,ASMC,ASMK,SPAO,CMTI
5	.514	.101	.040	.186	1.245	0.516	0.904	0.405	0.679	0.709	162	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMTI,CM2T
6	.513	.103	.038	.179	1.254	0.538	0.892	0.362	0.691	0.871	148	7	ASAR,ASWK,ASAS,ASMK,SPAO,CMTI,CM1T
7	.513	.105	.033	.174	1.289	0.426	0.959	0.290	0.582	0.630	156	7	ASAR,ASWK,ASAS,ASMK,SPAO,SPRS,CMTI
8	.513	.103	.036	.181	1.245	0.543	0.899	0.357	0.711	0.897	160	8	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMTI,CM1T
9	.513	.107	.035	.181	1.267	0.411	0.943	0.281	0.575	0.644	164	8	ASAR,ASWK,ASAS,ASMK,SPAO,SPRS,CMST,CMTI
10	.513	.103	.039	.179	1.262	0.510	0.897	0.409	0.682	0.728	148	7	ASAR,ASWK,ASAS,ASMK,SPAO,CMTI,CM2T
11	.513	.103	.037	.182	1.252	0.510	0.903	0.405	0.704	0.761	160	8	ASAR,ASWK,ASAS,ASEI,ASMK,SPAO,CMTI,CM2T
12	.513	.105	.039	.185	1.234	0.552	0.880	0.356	0.678	0.858	156	8	ASAR,ASWK,ASAS,ASMK,SPAO,CMST,CMTI,CM1T
13	.513	.105	.040	.186	1.239	0.498	0.883	0.402	0.672	0.711	156	8	ASAR,ASWK,ASAS,ASMK,SPAO,CMST,CMTI,CM2T
14	.513	.105	.037	.182	1.243	0.536	0.886	0.385	0.687	0.853	161	8	ASAR,ASWK,ASAS,ASMK,SPAO,SPOR,CMTI,CM1T
15	.513	.105	.034	.177	1.210	0.640	0.873	0.355	0.682	0.828	164	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CM1T
16	.512	.107	.035	.180	1.242	0.513	0.911	0.307	0.602	0.730	163	8	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMST,CMTI
17	.512	.103	.033	.171	1.230	0.643	0.881	0.354	0.700	0.860	156	7	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CM1T
18	.512	.106	.038	.183	1.251	0.538	0.890	0.429	0.681	0.705	161	8	ASAR,ASWK,ASAS,ASMK,SPAO,SPOR,CMTI,CM2T
19	.512	.104	.030	.165	1.232	0.585	0.864	0.378	0.711	0.910	164	7	ASAR,ASWK,ASAS,ASMC,ASMK,SPAO,CM1T
20	.512	.106	.033	.172	1.268	0.462	0.927	0.314	0.615	0.719	155	7	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,CMTI
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A8. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.501	.104	.046	.201	1.139	0.678	0.885	0.456	0.819	0.876	162	10	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
2	.502	.103	.046	.197	1.178	0.664	0.912	0.458	0.799	0.830	156	9	ASAR,ASWK,ASAS,ASGS,ASMK,CMST,CMTI,CM1T,CM2T
3	.502	.101	.046	.193	1.203	0.669	0.930	0.477	0.828	0.874	148	8	ASAR,ASWK,ASAS,ASGS,ASMK,CMTI,CM1T,CM2T
4	.496	.101	.045	.199	1.145	0.667	0.860	0.485	0.790	0.852	145	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,CMST,CMTI,CM1T,CM2T
5	.496	.098	.045	.196	1.155	0.662	0.882	0.465	0.836	0.845	135	9	ASAR,ASWK,ASAS,ASGS,ASNO,CMST,CMTI,CM1T,CM2T
6	.499	.103	.045	.195	1.200	0.696	0.927	0.476	0.828	0.942	158	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,CMTI,CM1T,CM2T
7	.494	.099	.045	.199	1.150	0.667	0.878	0.499	0.837	0.818	151	10	ASAR,ASWK,ASAS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
8	.500	.102	.044	.195	1.176	0.696	0.911	0.484	0.846	0.926	154	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMTI,CM1T,CM2T
9	.504	.104	.044	.196	1.191	0.670	0.913	0.508	0.802	0.874	161	9	ASAR,ASWK,ASAS,ASGS,ASMK,SPOR,CMTI,CM1T,CM2T
10	.495	.099	.044	.194	1.179	0.682	0.880	0.506	0.809	0.882	137	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,CMTI,CM1T,CM2T
11	.499	.101	.044	.194	1.197	0.681	0.924	0.526	0.836	0.861	164	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASPC,CMTI,CM1T,CM2T
12	.502	.104	.044	.194	1.151	0.589	0.895	0.454	0.804	0.751	155	9	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM2T
13	.494	.102	.044	.201	1.139	0.671	0.856	0.524	0.794	0.828	161	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
14	.504	.103	.044	.190	1.191	0.579	0.923	0.455	0.783	0.703	149	8	ASAR,ASWK,ASAS,ASGS,ASMK,CMST,CMTI,CM2T
15	.502	.104	.044	.193	1.177	0.595	0.916	0.452	0.791	0.773	159	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,CMST,CMTI,CM2T
16	.504	.093	.044	.201	1.193	0.649	0.946	0.408	0.735	0.756	150	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPRS,CMST,CMTI,CM1T,CM2T
17	.495	.099	.044	.193	1.161	0.580	0.887	0.494	0.816	0.813	144	9	ASAR,ASWK,ASAS,ASGS,ASNO,ASPC,CMST,CMTI,CM2T
18	.498	.101	.044	.193	1.156	0.580	0.869	0.482	0.772	0.761	138	9	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,CMST,CMTI,CM2T
19	.500	.106	.044	.195	1.150	0.610	0.875	0.462	0.832	0.935	148	9	ASAR,ASWK,ASAS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
20	.499	.107	.044	.194	1.180	0.596	0.898	0.461	0.820	0.962	152	9	ASAR,ASWK,ASAS,ASCS,ASMK,CMST,CMTI,CM1T,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A8. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.505	.096	.044	.204	1.187	0.652	0.931	0.420	0.723	0.744	163	11	ASAR,ASWK,ASAS,ASGS,ASNO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
2	.503	.097	.043	.203	1.183	0.654	0.926	0.417	0.712	0.793	160	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,SPRS,CMST,CMTI,CM1T,CM2T
3	.502	.094	.042	.202	1.198	0.646	0.954	0.407	0.747	0.766	162	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASNO,SPRS,CMST,CMTI,CM1T,CM2T
4	.501	.104	.046	.201	1.139	0.678	0.885	0.456	0.819	0.876	162	10	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,CMST,CMTI,CM1T,CM2T
5	.494	.102	.044	.201	1.139	0.671	0.856	0.524	0.794	0.828	161	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
6	.504	.093	.044	.201	1.193	0.649	0.946	0.408	0.735	0.756	150	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPRS,CMST,CMTI,CM1T,CM2T
7	.512	.096	.044	.201	1.218	0.560	0.909	0.348	0.648	0.666	164	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPAO,SPRS,CMST,CMTI,CM2T
8	.499	.103	.043	.201	1.147	0.678	0.859	0.505	0.771	0.848	158	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,SPOR,CMST,CMTI,CM1T,CM2T
9	.495	.102	.043	.200	1.152	0.663	0.871	0.479	0.804	0.866	157	11	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASNO,CMST,CMTI,CM1T,CM2T
10	.497	.102	.043	.200	1.150	0.678	0.871	0.525	0.806	0.814	164	11	ASAR,ASWK,ASAS,ASGS,ASNO,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
11	.508	.097	.043	.200	1.196	0.659	0.858	0.406	0.703	0.839	156	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPAO,CMST,CMTI,CM1T,CM2T
12	.505	.096	.043	.200	1.246	0.587	0.972	0.406	0.704	0.721	157	10	ASAR,ASWK,ASAS,ASGS,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
13	.492	.100	.042	.200	1.156	0.662	0.886	0.492	0.848	0.833	163	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
14	.496	.101	.045	.199	1.145	0.667	0.860	0.485	0.790	0.852	145	10	ASAR,ASWK,ASAS,ASCS,ASGS,ASNO,CMST,CMTI,CM1T,CM2T
15	.507	.098	.042	.199	1.222	0.569	0.887	0.355	0.671	0.806	157	10	ASAR,ASWK,ASAS,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
16	.501	.103	.041	.199	1.187	0.597	0.897	0.443	0.706	0.783	159	11	ASAR,ASWK,ASAS,ASCS,ASNO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
17	.509	.093	.042	.199	1.240	0.636	0.923	0.358	0.673	0.790	163	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPAO,SPRS,CMTI,CM1T,CM2T
18	.512	.095	.042	.199	1.216	0.644	0.909	0.349	0.654	0.792	164	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPAO,SPRS,CMST,CMTI,CM1T
19	.494	.099	.045	.199	1.150	0.667	0.878	0.499	0.837	0.818	151	10	ASAR,ASWK,ASAS,ASGS,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
20	.506	.097	.043	.199	1.196	0.568	0.939	0.419	0.710	0.620	156	10	ASAR,ASWK,ASAS,ASGS,ASNO,SPOR,SPRS,CMST,CMTI,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A8. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.438	.129	.011	.139	0.906	0.683	0.684	0.478	0.310	0.318	140	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.451	.123	.019	.156	0.939	0.662	0.697	0.488	0.549	0.642	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
3	.453	.121	.021	.154	0.947	0.662	0.722	0.491	0.589	0.670	137	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T
4	.450	.132	.010	.146	0.952	0.673	0.708	0.467	0.368	0.311	153	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST
5	.451	.122	.020	.163	0.954	0.664	0.705	0.489	0.572	0.630	154	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
6	.453	.130	.011	.140	0.957	0.684	0.714	0.486	0.364	0.300	137	7	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
7	.454	.121	.021	.159	0.960	0.675	0.711	0.507	0.568	0.657	138	8	ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
8	.452	.130	.013	.143	0.960	0.663	0.731	0.466	0.388	0.317	143	7	ASAR,ASWK,ASMK,ASNO,ASPC,SPOR,CMST
9	.452	.121	.020	.156	0.961	0.594	0.711	0.433	0.536	0.586	147	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
10	.452	.120	.023	.162	0.962	0.662	0.728	0.488	0.610	0.663	144	8	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
11	.454	.119	.023	.155	0.970	0.590	0.736	0.435	0.570	0.649	137	7	ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM2T
12	.450	.119	.018	.148	0.971	0.665	0.714	0.497	0.570	0.694	139	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
13	.459	.125	.018	.160	0.972	0.657	0.715	0.480	0.551	0.627	160	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST,CM1T
14	.454	.120	.014	.149	0.975	0.631	0.781	0.380	0.402	0.403	154	8	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,CMST
15	.461	.123	.020	.158	0.976	0.652	0.732	0.478	0.579	0.657	150	8	ASAR,ASWK,ASMK,ASNO,ASPC,SPOR,CMST,CM1T
16	.450	.127	.018	.157	0.977	0.504	0.761	0.386	0.381	0.357	146	8	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI
17	.461	.123	.017	.154	0.979	0.669	0.721	0.501	0.546	0.660	144	8	ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CM1T
18	.457	.120	.015	.143	0.980	0.641	0.789	0.395	0.391	0.424	138	7	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,CMST
19	.458	.122	.026	.172	0.981	0.550	0.749	0.456	0.558	0.630	153	9	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CMTI,CM1T
20	.447	.131	.007	.135	0.983	0.655	0.724	0.465	0.380	0.316	145	7	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A8. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--							
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test	
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV Predictor Variables
1	.438	.129	.011	.139	0.906	0.683	0.684	0.478	0.310	0.318	140	7 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST
2	.451	.123	.019	.156	0.939	0.662	0.697	0.488	0.549	0.642	147	8 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T
3	.451	.122	.020	.163	0.954	0.664	0.705	0.489	0.572	0.630	154	9 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM1T,CM2T
4	.467	.117	.015	.146	1.033	0.636	0.705	0.427	0.295	0.275	134	7 ASAR,ASWK,ASCS,ASNO,ASPC,SPAO,CMST
5	.465	.123	.016	.152	1.024	0.598	0.708	0.429	0.289	0.315	161	8 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,CMST
6	.450	.132	.010	.146	0.952	0.673	0.708	0.467	0.368	0.311	153	8 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST
7	.452	.121	.020	.156	0.961	0.594	0.711	0.433	0.536	0.586	147	8 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CMST,CM2T
8	.454	.121	.021	.159	0.960	0.675	0.711	0.507	0.568	0.657	138	8 ASAR,ASWK,ASCS,ASMK,ASNO,CMST,CM1T,CM2T
9	.473	.112	.020	.158	1.048	0.640	0.712	0.449	0.465	0.586	141	8 ASAR,ASWK,ASCS,ASNO,ASPC,SPAO,CMST,CM1T
10	.453	.130	.011	.140	0.957	0.684	0.714	0.486	0.364	0.300	137	7 ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST
11	.468	.122	.016	.147	1.030	0.609	0.714	0.448	0.284	0.308	145	7 ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST
12	.450	.119	.018	.148	0.971	0.665	0.714	0.497	0.570	0.694	139	7 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,CM1T
13	.459	.125	.018	.160	0.972	0.657	0.715	0.480	0.551	0.627	160	9 ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPOR,CMST,CM1T
14	.470	.113	.018	.160	1.056	0.636	0.717	0.447	0.482	0.579	148	9 ASAR,ASWK,ASCS,ASNO,ASPC,SPAO,CMST,CM1T,CM2T
15	.458	.119	.016	.158	0.993	0.679	0.717	0.474	0.566	0.566	140	9 ASAR,ASWK,ASCS,ASNO,ASPC,SPOR,CMST,CM1T,CM2T
16	.468	.121	.012	.148	1.043	0.634	0.717	0.429	0.331	0.266	147	8 ASAR,ASWK,ASCS,ASNO,ASPC,SPAO,SPOR,CMST
17	.475	.117	.021	.160	1.044	0.618	0.721	0.470	0.461	0.667	152	8 ASAR,ASWK,ASCS,ASMK,ASNO,SPAO,CMST,CM1T
18	.473	.115	.017	.159	1.055	0.636	0.721	0.451	0.474	0.554	154	9 ASAR,ASWK,ASCS,ASNO,ASPC,SPAO,SPOR,CMST,CM1T
19	.461	.123	.017	.154	0.979	0.669	0.721	0.501	0.546	0.660	144	8 ASAR,ASWK,ASCS,ASMK,ASNO,SPOR,CMST,CM1T
20	.453	.121	.021	.154	0.947	0.662	0.722	0.491	0.589	0.670	137	7 ASAR,ASWK,ASMK,ASNO,ASPC,CMST,CM1T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2 ASAR,ASWK

(Continued)

Table A8. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 134-164 Min. Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.465	.118	.017	.143	1.189	0.490	0.872	0.318	0.144	0.092	144	6	ASAR,ASWK,ASMK,SPAO,SPRS,CMST
2	.444	.120	.013	.135	1.149	0.600	0.941	0.378	0.149	0.280	139	6	ASAR,ASWK,ASMK,ASPC,SPRS,CMST
3	.466	.117	.013	.129	1.204	0.448	0.883	0.291	0.151	0.092	136	5	ASAR,ASWK,ASMK,SPAO,SPRS
4	.462	.119	.015	.147	1.179	0.494	0.861	0.297	0.158	0.236	160	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMST
5	.463	.117	.012	.134	1.195	0.450	0.873	0.276	0.166	0.204	152	6	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS
6	.457	.124	.013	.138	1.168	0.507	0.800	0.306	0.175	0.255	145	6	ASAR,ASWK,ASMK,ASPC,SPAO,CMST
7	.457	.122	.009	.123	1.189	0.453	0.812	0.290	0.187	0.212	137	5	ASAR,ASWK,ASMK,ASPC,SPAO
8	.444	.120	.015	.145	1.107	0.656	0.926	0.388	0.188	0.190	149	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS,CMST
9	.466	.118	.019	.152	1.145	0.576	0.854	0.308	0.193	0.211	154	7	ASAR,ASWK,ASCS,ASMK,SPAO,SPRS,CMST
10	.442	.118	.011	.132	1.140	0.647	0.949	0.372	0.196	0.204	141	6	ASAR,ASWK,ASCS,ASMK,ASPC,SPRS
11	.472	.115	.028	.166	1.214	0.460	0.910	0.375	0.199	0.205	150	7	ASAR,ASWK,ASMK,SPAO,SPRS,CMST,CMTI
12	.465	.116	.015	.137	1.169	0.540	0.870	0.280	0.199	0.232	146	6	ASAR,ASWK,ASCS,ASMK,SPAO,SPRS
13	.473	.113	.025	.157	1.224	0.460	0.919	0.370	0.200	0.239	142	6	ASAR,ASWK,ASMK,SPAO,SPRS,CMTI
14	.468	.123	.015	.147	1.171	0.473	0.854	0.315	0.201	0.112	157	7	ASAR,ASWK,ASMK,SPAO,SPOR,SPRS,CMST
15	.460	.112	.017	.150	1.197	0.509	0.870	0.248	0.201	0.210	143	7	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS,CMST
16	.462	.117	.014	.143	1.162	0.534	0.863	0.258	0.206	0.214	162	7	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPRS
17	.459	.109	.011	.133	1.228	0.488	0.889	0.236	0.208	0.220	135	6	ASAR,ASWK,ASCS,ASPC,SPAO,SPRS
18	.469	.121	.012	.134	1.184	0.431	0.864	0.288	0.209	0.084	149	6	ASAR,ASWK,ASMK,SPAO,SPOR,SPRS
19	.461	.122	.018	.144	1.130	0.582	0.793	0.346	0.211	0.293	139	6	ASAR,ASWK,ASCS,ASMK,SPAO,CMST
20	.470	.114	.022	.158	1.217	0.483	0.911	0.368	0.214	0.171	158	7	ASAR,ASWK,ASMK,ASPC,SPAO,SPRS,CMTI
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9.

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.515	.099	.038	.193	1.234	0.628	0.926	0.346	0.659	0.820	199	10	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMTI,CM1T
2	.515	.100	.038	.193	1.240	0.519	0.929	0.371	0.655	0.688	199	10	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMTI,CM2T
3	.514	.102	.040	.200	1.219	0.540	0.913	0.364	0.648	0.679	207	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMST,CMTI,CM2T
4	.514	.101	.039	.198	1.215	0.635	0.912	0.358	0.650	0.813	207	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMST,CMTI,CM1T
5	.514	.104	.034	.189	1.237	0.512	0.933	0.289	0.597	0.653	200	10	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMST,CMTI
6	.514	.103	.042	.202	1.220	0.538	0.918	0.369	0.633	0.604	198	11	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
7	.514	.103	.040	.200	1.216	0.634	0.917	0.330	0.637	0.748	198	11	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
8	.513	.101	.039	.199	1.218	0.631	0.926	0.332	0.657	0.793	197	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM1T
9	.513	.101	.038	.196	1.227	0.624	0.919	0.336	0.656	0.797	212	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPOR,SPRS,CMTI,CM1T
10	.513	.102	.040	.200	1.222	0.534	0.928	0.359	0.652	0.658	197	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM2T
11	.513	.099	.041	.201	1.227	0.631	0.920	0.365	0.668	0.816	206	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPRS,CMTI,CM1T,CM2T
12	.513	.103	.036	.191	1.240	0.544	0.912	0.337	0.661	0.824	198	10	ASAR,ASWK,ASAS,ASMC,ASMK,SPAO,SPOR,SPRS,CMTI,CM1T
13	.513	.102	.039	.198	1.232	0.524	0.921	0.369	0.653	0.660	212	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPOR,SPRS,CMTI,CM2T
14	.513	.102	.037	.191	1.219	0.636	0.887	0.377	0.689	0.845	197	10	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPOR,CMTI,CM1T
15	.513	.104	.038	.193	1.244	0.521	0.914	0.374	0.658	0.685	198	10	ASAR,ASWK,ASAS,ASMC,ASMK,SPAO,SPOR,SPRS,CMTI,CM2T
16	.513	.101	.034	.187	1.242	0.550	0.925	0.337	0.681	0.870	197	10	ASAR,ASWK,ASAS,ASEI,ASMC,ASMK,SPAO,SPRS,CMTI,CM1T
17	.513	.101	.042	.203	1.227	0.619	0.923	0.377	0.655	0.750	197	11	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
18	.513	.102	.038	.196	1.229	0.616	0.931	0.332	0.663	0.776	202	11	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,SPAO,SPOR,SPRS,CMTI,CM1T
19	.513	.102	.041	.201	1.207	0.670	0.916	0.328	0.651	0.841	195	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM1T
20	.513	.105	.033	.186	1.250	0.479	0.941	0.300	0.606	0.601	205	10	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,SPAO,SPOR,SPRS,CMTI
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY DISCRIMINANT VALIDITY ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.512	.102	.044	.209	1.190	0.667	0.900	0.347	0.662	0.821	198	12	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
2	.511	.102	.044	.209	1.199	0.669	0.909	0.347	0.662	0.831	202	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
3	.506	.103	.044	.207	1.182	0.681	0.940	0.429	0.714	0.795	194	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
4	.500	.107	.044	.205	1.149	0.698	0.886	0.530	0.790	0.870	195	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASPC,SPOR,CMST,CMTI,CM1T,CM2T
5	.510	.104	.043	.211	1.186	0.663	0.895	0.361	0.657	0.794	211	13	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
6	.512	.103	.043	.208	1.210	0.631	0.911	0.367	0.646	0.738	205	12	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
7	.510	.104	.043	.211	1.193	0.664	0.902	0.359	0.658	0.802	215	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
8	.504	.106	.043	.205	1.154	0.681	0.884	0.464	0.779	0.897	201	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,SPOR,CMST,CMTI,CM1T,CM2T
9	.509	.104	.043	.206	1.170	0.683	0.854	0.435	0.705	0.847	199	12	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
10	.508	.104	.043	.206	1.183	0.681	0.865	0.442	0.702	0.861	203	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
11	.497	.106	.043	.204	1.153	0.682	0.901	0.495	0.827	0.885	194	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
12	.509	.102	.043	.210	1.183	0.670	0.894	0.380	0.668	0.797	214	13	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
13	.503	.103	.043	.201	1.162	0.663	0.892	0.491	0.801	0.821	194	11	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
14	.501	.104	.043	.204	1.155	0.679	0.888	0.481	0.801	0.879	204	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,ASPC,CMST,CMTI,CM1T,CM2T
15	.510	.105	.042	.206	1.183	0.675	0.867	0.418	0.692	0.863	200	12	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
16	.500	.104	.042	.204	1.145	0.678	0.880	0.477	0.805	0.854	200	12	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,ASNO,ASPC,CMST,CMTI,CM1T,CM2T
17	.509	.105	.042	.206	1.186	0.671	0.874	0.396	0.710	0.896	199	12	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,SPAO,CMST,CMTI,CM1T,CM2T
18	.513	.101	.042	.203	1.227	0.619	0.923	0.377	0.655	0.750	197	11	ASAR,ASWK,ASAS,ASGS,ASMK,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
19	.512	.104	.042	.206	1.194	0.579	0.901	0.363	0.645	0.661	204	12	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
20	.513	.102	.042	.202	1.210	0.581	0.917	0.346	0.647	0.705	195	11	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY BROGDEN INDEX ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.509	.105	.041	.212	1.187	0.659	0.901	0.350	0.668	0.804	223	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
2	.507	.105	.041	.211	1.183	0.663	0.892	0.362	0.659	0.810	221	14	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
3	.510	.104	.043	.211	1.186	0.663	0.895	0.361	0.657	0.794	211	13	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
4	.510	.104	.043	.211	1.193	0.664	0.902	0.359	0.658	0.802	215	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
5	.507	.104	.040	.210	1.190	0.666	0.904	0.340	0.675	0.849	220	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
6	.509	.102	.043	.210	1.183	0.670	0.894	0.380	0.668	0.797	214	13	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
7	.510	.103	.042	.210	1.192	0.664	0.906	0.342	0.672	0.830	210	13	ASAR,ASWK,ASAS,ASEI,ASGS,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
8	.506	.104	.040	.210	1.182	0.671	0.891	0.376	0.669	0.814	224	14	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
9	.509	.103	.042	.210	1.188	0.668	0.896	0.345	0.663	0.838	208	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
10	.510	.103	.042	.210	1.196	0.662	0.911	0.344	0.676	0.845	214	13	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASMK,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
11	.505	.101	.039	.209	1.200	0.637	0.884	0.354	0.655	0.822	216	14	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
12	.510	.102	.041	.209	1.192	0.665	0.896	0.352	0.663	0.846	220	13	ASAR,ASWK,ASAS,ASGS,ASMC,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
13	.512	.102	.044	.209	1.190	0.667	0.900	0.347	0.662	0.821	198	12	ASAR,ASWK,ASAS,ASGS,ASMK,ASNO,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
14	.504	.099	.039	.209	1.203	0.637	0.890	0.373	0.666	0.765	222	14	ASAR,ASWK,ASAS,ASGS,ASMC,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
15	.502	.102	.037	.209	1.197	0.645	0.890	0.393	0.664	0.773	222	15	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
16	.504	.099	.039	.209	1.203	0.640	0.900	0.383	0.673	0.745	212	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
17	.509	.103	.042	.209	1.193	0.672	0.903	0.390	0.667	0.808	218	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMK,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
18	.504	.101	.039	.209	1.202	0.643	0.894	0.365	0.661	0.799	206	14	ASAR,ASWK,ASAS,ASCS,ASEI,ASGS,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
19	.504	.099	.038	.209	1.209	0.630	0.900	0.338	0.672	0.796	218	14	ASAR,ASWK,ASAS,ASEI,ASGS,ASMC,ASNO,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
20	.510	.102	.041	.209	1.194	0.667	0.898	0.341	0.669	0.866	224	13	ASAR,ASWK,ASAS,ASCS,ASGS,ASMC,ASMK,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY WHITE-BLACK SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--S U B G R O U P	D I F F E R E N C E--							
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.474	.112	.020	.177	1.059	0.589	0.853	0.368	0.549	0.630	196	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CM1T,CM2T
2	.473	.119	.018	.171	1.061	0.578	0.767	0.402	0.445	0.558	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
3	.470	.120	.016	.173	1.067	0.572	0.770	0.381	0.461	0.545	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
4	.473	.120	.024	.184	1.070	0.570	0.757	0.434	0.492	0.593	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
5	.477	.110	.028	.192	1.071	0.591	0.871	0.442	0.545	0.618	202	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
6	.479	.111	.027	.186	1.072	0.591	0.874	0.404	0.532	0.626	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM1T
7	.472	.119	.017	.168	1.074	0.500	0.775	0.323	0.435	0.416	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
8	.482	.115	.021	.174	1.074	0.610	0.784	0.389	0.520	0.696	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T
9	.479	.116	.019	.176	1.078	0.602	0.786	0.387	0.533	0.686	202	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
10	.479	.110	.028	.187	1.080	0.560	0.879	0.441	0.528	0.493	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPOR,SPRS,CMST,CMTI,CM2T
11	.475	.116	.027	.188	1.082	0.536	0.797	0.391	0.457	0.588	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
12	.477	.117	.025	.185	1.082	0.557	0.797	0.373	0.448	0.558	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
13	.471	.117	.014	.165	1.084	0.564	0.782	0.372	0.465	0.573	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CM1T,CM2T
14	.474	.118	.025	.189	1.084	0.546	0.797	0.400	0.461	0.544	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
15	.481	.113	.029	.191	1.084	0.609	0.800	0.416	0.527	0.709	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
16	.481	.115	.019	.170	1.086	0.529	0.792	0.350	0.512	0.555	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM2T
17	.484	.113	.028	.189	1.087	0.627	0.807	0.436	0.535	0.679	198	12	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
18	.484	.114	.027	.187	1.087	0.624	0.804	0.387	0.518	0.694	201	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T
19	.477	.116	.028	.189	1.087	0.558	0.806	0.385	0.468	0.539	199	12	ASAR,ASWK,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
20	.481	.115	.026	.191	1.087	0.617	0.804	0.422	0.530	0.684	208	13	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY WHITE-HISPANIC SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV	Predictor Variables
1	.473	.120	.024	.184	1.070	0.570	0.757	0.434	0.492	0.593	194	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI,CM1T,CM2T
2	.473	.119	.018	.171	1.061	0.578	0.767	0.402	0.445	0.558	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T
3	.470	.120	.016	.173	1.067	0.572	0.770	0.381	0.461	0.545	203	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
4	.472	.119	.017	.168	1.074	0.500	0.775	0.323	0.435	0.416	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
5	.471	.117	.014	.165	1.084	0.564	0.782	0.372	0.465	0.573	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CM1T,CM2T
6	.482	.115	.021	.174	1.074	0.610	0.784	0.389	0.520	0.696	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T
7	.479	.116	.019	.176	1.078	0.602	0.786	0.387	0.533	0.686	202	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
8	.486	.114	.023	.180	1.123	0.598	0.790	0.357	0.580	0.776	210	12	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
9	.488	.113	.023	.176	1.123	0.603	0.790	0.364	0.570	0.794	203	11	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T
10	.489	.112	.024	.178	1.125	0.599	0.791	0.375	0.578	0.753	200	11	ASAR,ASWK,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM1T,CM2T
11	.481	.115	.019	.170	1.086	0.529	0.792	0.350	0.512	0.555	195	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM2T
12	.487	.112	.024	.178	1.129	0.597	0.793	0.367	0.583	0.807	197	11	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,CMST,CM1T,CM2T
13	.488	.114	.022	.174	1.131	0.526	0.796	0.301	0.568	0.652	203	11	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CM2T
14	.474	.118	.025	.189	1.084	0.546	0.797	0.400	0.461	0.544	209	13	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
15	.477	.117	.025	.185	1.082	0.557	0.797	0.373	0.448	0.558	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
16	.475	.116	.027	.188	1.082	0.536	0.797	0.391	0.457	0.588	196	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI,CM1T,CM2T
17	.488	.113	.023	.176	1.131	0.599	0.798	0.376	0.575	0.803	194	11	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,SPAO,SPOR,CMST,CM1T,CM2T
18	.479	.114	.017	.168	1.101	0.601	0.799	0.390	0.541	0.729	194	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CM1T,CM2T
19	.481	.113	.029	.191	1.084	0.609	0.800	0.416	0.527	0.709	195	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,CMST,CMTI,CM1T,CM2T
20	.488	.116	.020	.167	1.134	0.553	0.801	0.316	0.490	0.555	196	10	ASAR,ASWK,ASCS,ASMC,ASMK,ASNO,ASPC,SPAO,SPOR,CMST
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

(Continued)

Table A9. (Cont'd)

On-the-Job Hands-on Test Performance: Validities of Top 20 Equations Ranked According to Each Performance Index, at the 194-224 Interval

TOP 20 EQUATIONS RANKED BY MALE-FEMALE SUBGROUP DIFFERENCE ESTIMATE (SUMMARY ACROSS 9 MOS):

					--SUBGROUP DIFFERENCE--								
	Validty	Est.	Discr	Brog.	Whte-Blck		Whte-Hisp		Male-Fem.		Test		
Rank	Mean	S.D.	Vldty	Index	Index	Mean	Rnge	Mean	Rnge	Mean	Rnge	Time	NV Predictor Variables
1	.474	.121	.022	.177	1.087	0.498	0.810	0.329	0.338	0.259	195	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
2	.481	.114	.019	.168	1.131	0.597	0.879	0.274	0.344	0.276	197	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST
3	.483	.112	.026	.182	1.151	0.563	0.903	0.351	0.363	0.307	203	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
4	.484	.111	.023	.172	1.173	0.510	0.919	0.344	0.368	0.331	195	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI
5	.483	.112	.025	.181	1.111	0.518	0.864	0.316	0.368	0.395	196	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPRS,CMST,CMTI
6	.480	.115	.016	.169	1.091	0.556	0.836	0.301	0.371	0.316	203	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST
7	.479	.114	.012	.158	1.115	0.533	0.851	0.272	0.376	0.346	195	10	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS
8	.483	.114	.023	.183	1.111	0.530	0.861	0.333	0.390	0.348	209	12	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
9	.486	.112	.026	.183	1.112	0.539	0.867	0.298	0.392	0.335	199	11	ASAR,ASWK,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI
10	.483	.113	.020	.173	1.135	0.485	0.877	0.324	0.395	0.373	201	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMTI
11	.474	.116	.028	.183	1.152	0.583	0.857	0.457	0.409	0.376	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
12	.475	.116	.028	.184	1.140	0.590	0.850	0.419	0.417	0.525	196	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T
13	.481	.118	.022	.177	1.098	0.555	0.819	0.347	0.418	0.394	194	11	ASAR,ASWK,ASCS,ASGS,ASMK,ASNO,ASPC,SPAO,SPOR,CMST,CMTI
14	.468	.119	.019	.172	1.127	0.615	0.823	0.372	0.432	0.508	197	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CM1T,CM2T
15	.472	.117	.027	.187	1.142	0.583	0.849	0.450	0.432	0.504	203	12	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM1T,CM2T
16	.473	.115	.025	.180	1.155	0.583	0.859	0.441	0.432	0.515	195	11	ASAR,ASWK,ASCS,ASMK,ASPC,SPAO,SPOR,SPRS,CMTI,CM1T,CM2T
17	.472	.119	.017	.168	1.074	0.500	0.775	0.323	0.435	0.416	196	11	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
18	.484	.109	.021	.169	1.162	0.539	0.882	0.354	0.439	0.391	194	10	ASAR,ASWK,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CM2T
19	.486	.107	.029	.185	1.169	0.565	0.896	0.448	0.439	0.391	200	11	ASAR,ASWK,ASGS,ASMK,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
20	.476	.117	.025	.185	1.092	0.490	0.803	0.403	0.441	0.423	202	12	ASAR,ASWK,ASCS,ASMK,ASNO,ASPC,SPAO,SPOR,SPRS,CMST,CMTI,CM2T
BASE	.411	.123	.003	.081	1.246	0.509	0.917	0.365	0.219	0.201	73	2	ASAR,ASWK

Appendix B

LVII SOLDIERS MEETING PREDICTOR/CRITERION
SETWISE DELECTION DATA REQURIEMENTS FOR
VALIDATION OF EXPERIMENTAL BATTERY PREDICTOR COMPOSITES
AGAINST PERFORMANCE CRITERIA

Table B.1

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Core Technical Proficiency Factor by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	333	322	112	301	297	309
13B	170	165	152	159	148	156
19K	156	130	130	122	123	129
63B	169	147	139	136	132	140
71L	147	115	104	105	109	102
88M[a]	84	56	54	51	52	53
91A	205	191	174	185	165	183
95B	160	149	140	142	133	145
Total	1,424	1,275	1,005	1,201	1,159	1,217

[a] MOS 88M not included in LVII validity analyses.

Table B.2

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against General Soldiering Proficiency Factor by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B[a]	--	--	--	--	--	--
13B	170	165	152	159	148	156
19K	156	130	130	122	123	129
63B	169	147	139	136	132	140
71L	147	115	104	105	109	102
88M[b]	84	56	54	51	52	53
91A	205	191	174	185	165	183
95B	160	149	140	142	133	145
Total	1,091	953	893	900	862	908

[a] MOS 11B not included in LVII validity analyses for General Soldiering Proficiency.
[b] MOS 88M not included in LVII validity analyses.

Table B.3

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Achievement and Effort Factor by MOS

MOS	Predictor Sets					
	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	310	300	105	280	276	286
13B	157	152	140	143	136	143
19K	138	116	115	110	112	115
63B	183	157	148	145	140	149
71L	140	109	99	100	103	96
88M[a]	80	52	50	48	49	50
91A	179	165	151	158	144	157
95B	146	137	129	130	121	133
Total	1,333	1,188	937	1,114	1,081	1,129

[a] MOS 88M not included in LVII validity analyses.

Table B.4

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Personal Discipline Factor by MOS

MOS	Predictor Sets					
	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	320	310	109	290	286	296
13B	169	164	152	155	147	154
19K	143	120	120	113	116	119
63B	184	158	149	146	141	150
71L	144	111	101	102	105	98
88M[a]	83	55	53	50	52	52
91A	205	190	173	182	164	181
95B	161	150	141	143	134	146
Total	1,409	1,258	998	1,181	1,145	1,196

[a] MOS 88M not included in LVII validity analyses.

Table B.5

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Physical Fitness and Military Bearing Factor by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	309	299	107	281	278	286
13B	162	157	145	148	141	147
19K	137	117	117	110	113	116
63B	177	151	142	140	134	143
71L	143	110	100	101	104	97
88M[a]	81	54	52	49	51	51
91A	194	179	164	172	154	171
95B	158	147	138	141	131	143
Total	1,361	1,214	965	1,142	1,106	1,154

[a] MOS 88M not included in LVII validity analyses.

Table B.6

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Leadership Factor by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	298	290	103	270	266	277
13B	131	127	117	122	110	119
19K	114	99	100	93	98	98
63B	169	146	137	136	130	139
71L	130	104	95	96	99	94
88M[a]	69	48	45	44	45	46
91A	183	170	158	165	153	164
95B	146	136	126	130	122	132
Total	1,240	1,120	881	1,056	1,023	1,069

[a] MOS 88M not included in LVII validity analyses.

Table B.7

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Hands-On Total Score by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	333	322	112	301	297	309
13B	171	166	153	160	148	157
19K	156	130	130	122	123	129
63B	171	149	141	138	134	142
71L	149	116	105	106	110	103
88M[a]	84	56	54	51	52	53
91A	206	192	175	186	166	184
95B	160	149	140	142	133	145
Total	1,430	1,280	1,010	1,206	1,163	1,222

[a] MOS 88M not included in LVII validity analyses.

Table B.8

Soldiers in LVII Meeting Predictor/Criterion Setwise Deletion Data Requirements for Validation of ASVAB Operational Scores and Spatial, Computer, JOB, ABLE, and AVOICE Experimental Battery Predictor Composites Against Job Knowledge Total Score by MOS

	Predictor Sets					
MOS	ASVAB	Spatial	Computer	JOB	ABLE	AVOICE
11B	341	330	115	309	305	316
13B	178	173	160	164	156	163
19K	164	138	137	130	131	137
63B	186	160	151	148	143	152
71L	149	116	105	106	110	103
88M[a]	85	57	55	52	53	54
91A	216	201	183	193	173	192
95B	164	153	144	146	137	149
Total	1,483	1,328	1,050	1,248	1,208	1,266

[a] MOS 88M not included in LVII validity analyses.

Appendix C

ARMY JOB SATISFACTION QUESTIONNAIRE (AJSQ)

LVII Version

Note the following changes from the LVI/CVII version of the AJSQ:

Section I

- *Items 13, 16, and 23 were added.*

Section II

- *The following four items were dropped:*
    - * *When you decided to enlist, how sure were you about what training you wanted?*
    - * *In high school, what education program were you in?*
    - * *Has your father, mother, or guardian ever served, or is (s)he now serving in the military?*
    - * *In what MOS do you now work (i.e., your current duty assignment)?*
- *Items 33 through 57 were added.*

Name ____________________

SS# ____ ____ ____

Army Job Satisfaction Questionnaire

Section I:

This section asks questions about your supervisors, your co-workers, and your job/career. We would like to know how you feel about each of these aspects of Army life. Please read each statement carefully and decide how you feel about it. Mark the box that corresponds to your level of satisfaction with the aspect of Army life described by the statement. For example, mark "Very Satisfied" if it is much better than you hoped it would be. Or, you might mark "Very Dissatisfied" if it is much poorer than you would like it to be.

Mark your answers by putting an "X" over the appropriate box like this: ☒

	Very Dissatisfied	Dissatisfied	Neither Satisfied nor Dissatisfied	Satisfied	Very Satisfied
A. Supervision: How satisfied are you with...					
1. Your immediate supervisor's capabilities?	☐	☐	☐	☐	☐
2. The way your immediate supervisor handles soldiers in your unit?	☐	☐	☐	☐	☐
3. Overall, how satisfied are you with the quality of supervision you are currently receiving?	☐	☐	☐	☐	☐
B. Co-Workers: How satisfied are you with...					
4. The amount of respect you get from the soldiers in your unit?	☐	☐	☐	☐	☐
5. The way soldiers in your unit work together to finish a job?	☐	☐	☐	☐	☐
6. Overall, how satisfied are you with your co-workers in the Army?	☐	☐	☐	☐	☐
C. Promotions: How satisfied are you with...					
7. Your chances for promotions in your current duty MOS?	☐	☐	☐	☐	☐
8. Your chances of getting ahead in the Army?	☐	☐	☐	☐	☐
9. Overall, how satisfied are you with your opportunities for promotion?	☐	☐	☐	☐	☐
D. Pay: How satisfied are you with the way your Army pay...					
10. Meets your personal and family needs?	☐	☐	☐	☐	☐
11. Covers the cost of living where you are currently assigned?	☐	☐	☐	☐	☐
12. Overall, how satisfied are you with your salary?	☐	☐	☐	☐	☐
E. Work (Your current duty assignment): How satisfied are you with ...					
13. The type of work you do?	☐	☐	☐	☐	☐
14. Your opportunity to do important things on the job?	☐	☐	☐	☐	☐
15. Your opportunity to use your abilities, experience, and training?	☐	☐	☐	☐	☐
16. Your opportunity to do interesting work?	☐	☐	☐	☐	☐
17. The amount of challenge in your work?	☐	☐	☐	☐	☐
18. Overall, how satisfied are you with your current duty assignment?	☐	☐	☐	☐	☐

Survey Approval Authority: U.S. Army Personnel Integration Command
Survey Control Number: ATNC-AO-91-52
RCS: MILPC-3

F. The Army as an organization: How satisfied are you with ...

	Very Dissatisfia	Dissatisfied	Neither Satisfiec nor Dissatisfiec	Satisfied	Very Satisfied
19. The way the Army treats its soldiers?	☐	☐	☐	☐	☐
20. Your opportunity to spend time with your family?	☐	☐	☐	☐	☐
21. The amount of leisure time you have?	☐	☐	☐	☐	☐
22. In general, how satisfied are you with all aspects of Army life?	☐	☐	☐	☐	☐

23. **All things considered, how do you feel about your job situation most of the time?**

- ☐ Best imaginable
- ☐ Highly favorable
- ☐ Satisfactory
- ☐ Neutral
- ☐ Poor
- ☐ Very bad
- ☐ Worst imaginable

Section II:

The questions in this section are related to your reasons for enlisting, your future plans, and the factors that may be influencing these plans. Please read each question carefully and answer as best you can. Mark your answers by putting an "X" over the appropriate box.

Mark the box indicating the importance of each reason for enlisting:

	Not at all important	Not very important	Somewhat important	Very important	Extremely important
24. To serve my country	☐	☐	☐	☐	☐
25. To get trained in a skill	☐	☐	☐	☐	☐
26. Money for education or training (e.g., college or vocational, technical, or business school)	☐	☐	☐	☐	☐
27. To be a soldier	☐	☐	☐	☐	☐
28. The physical training and challenge	☐	☐	☐	☐	☐
29. To learn to be a responsible mature person	☐	☐	☐	☐	☐

30. Would you prefer working in a different MOS?

- ☐ No
- ☐ Yes
- ☐ Not sure

31. If you could reenlist tomorrow, which MOS would you select?

- ☐ My Primary (AIT completed)MOS
- ☐ My Secondary (AIT completed)MOS
- ☐ A Duty MOS other than my Primary or Secondary

32. What do you think you will do after this enlistment?

- [ ] Leave the Army to find civilian employment
- [ ] Leave the Army to attend college
- [ ] Leave the Army for civilian vocational/ technical education
- [ ] Reenlist
- [ ] I do not know

33. What is the level of morale in you unit? If you work someplace else, answer this question for the place where you work.

- [ ] Very low
- [ ] Low
- [ ] Moderate
- [ ] High
- [ ] Very high

34. Overall, how fairly has the Army treated you?

- [ ] Very fairly
- [ ] Somewhat fairly
- [ ] Neither fairly nor unfairly
- [ ] Somewhat unfairly
- [ ] Very unfairly

	More interested	Neither more nor less interested	Less interested	Not sure
35. Do the recent developments in the Middle East (i.e., Operation Desert Shield/Storm) make you more or less interested in serving in the Army?	☐	☐	☐	☐
36. Do recent developments in Eastern Europe (e.g., the fall of the communist governments) make you more or less interested in serving in the Army?	☐	☐	☐	☐
37. Do likely reductions in the size of the U.S. Army make you more or less interested in serving in the Army?	☐	☐	☐	☐
38. Does the current state of the U.S. economy make you more or less interested in serving in the Army?	☐	☐	☐	☐

	Strong positive effect	Positive effect	No effect	Negative effect	Strong negative effect
39. What effect, if any, have your Army experiences had on the development of specific job knowledge, skills, and abilities that will help you obtain a civilian job?	☐	☐	☐	☐	☐
40. What effect, if any, have your Army experiences had on the development of personal characteristics and attitudes that will help you obtain a civilian job?	☐	☐	☐	☐	☐

Are you more concerned about these things today than you were two years ago?

	Much more concerned	More concerned	Neither more nor less concerned	Less concerned	Much less concerned
41. Your long-term opportunities in the Army	☐	☐	☐	☐	☐
42. The kind of work you might have to do in the Army	☐	☐	☐	☐	☐
43. The kind of work you will go into when you leave the Army	☐	☐	☐	☐	☐
44. Whether you would be able to get a civilian job quickly if you had to	☐	☐	☐	☐	☐
45. The financial burden on you and your family should you have to leave the Army unexpectedly	☐	☐	☐	☐	☐
46. Problems associated with deployment (e.g., separation from family, childcare arrangements, financial burdens)	☐	☐	☐	☐	☐

How much better or worse do you think these aspects would be for you in civilian life than they are for you in the Army?

	Much worse in civilian life	Worse in civilian life	About the same in both	Better in civilian life	Much better in civilian life	Don't know
47. Your income	☐	☐	☐	☐	☐	☐
48. Your ability to develop community ties	☐	☐	☐	☐	☐	☐
49. Support services for families	☐	☐	☐	☐	☐	☐
50. Quality of your co-workers	☐	☐	☐	☐	☐	☐
51. Personal freedom	☐	☐	☐	☐	☐	☐
52. Credit for doing a good job	☐	☐	☐	☐	☐	☐
53. Type of work you do	☐	☐	☐	☐	☐	☐
54. Your promotion opportunities	☐	☐	☐	☐	☐	☐
55. Your job security	☐	☐	☐	☐	☐	☐
56. Your treatment by supervisors	☐	☐	☐	☐	☐	☐
57. Your feelings of commitment to your employer	☐	☐	☐	☐	☐	☐